

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 2, 2002, 09:59:15 : Search time 72.0896 seconds
(without alignments)
297.593 Million cell updates/sec

Title: US-09-741-106-19

Perfect score: 905

Sequence: 1 DSEDEHTITDTELPPLK.....ECKNICDGPNGFQVDNYGT 161

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_101002.*

1:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21:	/SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22:	/SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23:	/SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	902	99.7	161	AAW00017	TFPI chimeric prot
2	900	99.4	276	AA922265	TFPI mutein, Lys36
3	900	99.4	276	AA92012	TFPI mutein K36R.
4	897	99.1	225	AAU02971	Angiotensin conver
5	897	99.1	276	AAW30311	Recombinant non-gl
6	897	99.1	276	AAW61535	Human tissue facto
7	897	99.1	304	AAW61535	Lipoprotein-associ
8	897	99.1	304	AAW61535	Tissue factor path
9	897	99.1	304	AAW61535	Human lipoprotein
10	897	99.1	352	AA92011	Ubiquitin-TFPI fus

11	892	98.6	160	13	AA23800	LACI fragment 1 -
12	891	98.5	304	10	AA92002	Human tissue facto
13	890	98.3	256	12	AA91169	Ser(Asp1-Thr255)-E
14	890	98.3	261	12	AA91171	Ser(Asp1-Glu245)-
15	890	98.3	262	12	AA91172	Ser(Asp1-Ser248)-
16	890	98.3	265	12	AA91170	Ser(Asp1-Ile253)-
17	886	97.9	304	21	AA970273	Human tissue facto
18	883	97.6	304	16	AA978389	Human lipoprotein-
19	881	97.3	277	14	AA97312	Non-glycosylated T
20	879	97.1	189	12	AA91167	(Asp1-Thr161)-EPI
21	876	96.8	304	21	AA970272	Human mutant tissu
22	836	92.4	304	14	AA942309	LACI gene product.
23	825	91.2	183	12	AA91146	Ser(Glu15-Thr161)
24	594	65.6	302	14	AA935001	LACI. Rattus ratt
25	594	65.6	302	17	AA98513	Lipoprotein-associ
26	425	47.0	110	12	AA91145	Ser-(Thr88-Thr161)
27	425	47.0	250	12	AA913675	Factor X-LACI hybr
28	408	45.1	71	13	AA923799	LACI fragment 90 -
29	377	41.7	128	20	AA982772	Bovine bikunin pep
30	371	41.0	128	20	AA982771	Bovine bikunin pep
31	370	40.9	128	20	AA982770	Bovine bikunin pep
32	347.5	38.4	111	14	AA93804	Synthetic yeast le
33	345	38.1	128	20	AA982768	Bovine bikunin pep
34	343	37.9	128	20	AA982769	Bovine bikunin pep
35	340	37.6	128	20	AA982766	Bovine bikunin pep
36	337	37.2	128	20	AA982767	Bovine bikunin pep
37	336	37.1	128	20	AA982785	Bovine bikunin pep
38	334	36.9	58	14	AA93802	TFPI 2. Homo sapi
39	334	36.9	58	16	AA981914	Lipoprotein-associ
40	334	36.9	58	16	AA978541	Human LACI-K2 doma
41	334	36.9	58	17	AA999189	Human aprotinin-11
42	334	36.9	58	19	AA964113	Human Kunitz-type
43	334	36.9	58	20	AA992860	US5880256 Seq id 3
44	334	36.9	83	17	AA999218	Human aprotinin-11
45	333	36.8	128	20	AA982787	Bovine bikunin pep

ALIGNMENTS

RESULT 1
AAW00017
ID AAW00017 standard; Protein; 161 AA.
AC AAW00017;
DT 30-SEP-1996 (first entry)
XX TFPI chimeric protein.
XX Kunitz-type domain; C-terminal tail; tissue factor pathway inhibitor;
KW TFPI; TFPI-2; cell surface localisation; glycosaminoglycan; heparin;
KW phospholipid; binding; chimeric protein; mutein; substitution;
KW PI-reactive site; sepsis; septic shock; thrombosis; up-regulation;
KW tissue factor; injury; trauma; endotoxin; TNF; cancer; IL-1;
KW tumour necrosis factor; interleukin.
XX Homo sapiens.
XX Key Location/Qualifiers
FH Domain 26..76
FT Note= "First Kunitz-type domain"
FT Misc-difference 36 /label= Lys36Arg
FT Domain 97..147 /note= "Second Kunitz-type domain"
FT Misc-difference 117 /label= Asn117Gln
XX WO9604378-A2.
XX 15-FEB-1996.
XX

XX DE TPPI mutein K36R.
 XX KW Tissue factor pathway inhibitor; TPPI; TPPI-2; Factor-VIIa;
 KW KW tissue factor; Factor-Xa; binding protein; tissue factor inhibitor;
 KW KW lipoprotein associated coagulation inhibitor;
 KW KW extrinsic pathway inhibitor; sepsis; septic shock;
 KW KW Saccharomyces cerevisiae.
 XX OS Synthetic.
 XX OS WO9604377-A1.
 XX PN 15-FEB-1996.
 XX PD
 XX PF
 XX XX 25-JUL-1995; 95WO-US09377.
 XX PF
 XX XX 05-AUG-1994; 94US-0286530.
 XX PR
 XX XX (CHIR) CHIRON CORP.
 XX PA
 XX PI Creasey AA, Innis MA;
 XX XX WPI; 1996-129393/13.
 XX DR
 XX XX Production of tissue factor pathway inhibitor in yeast cells - with
 PT isolation from the insoluble cell fraction, used to treat or prevent
 PT sepsis or septic shock
 XX PT
 XX PS Disclosure; Page 9; 36pp; English.
 XX PS
 CC A mutein (AAR92012) of tissue factor pathway inhibitor (TFPI)
 CC is prep'd. by site-directed mutagenesis of an encoding sequence.
 CC In the mutein, the lysine residue in the PI reactive site of the
 CC first kunitz-type domain of TPPI is replaced by arginine. This
 CC domain is required for the inhibition of Factor-VIIa/tissue
 CC factor (TF) complex. The mutein may be expressed in Saccharomyces
 CC cerevisiae transformants, esp. as a ubiquitin fusion protein,
 CC and as a Factor-VIIa/TF/Xa binding protein.
 XX CC
 XX SQ Sequence 276 AA;
 Query Match 99.4%; Score 900; DB 17; Length 276;
 Best Local Similarity 99.4%; Pred. No. 1.2e-82;
 Matches 160; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 DSEDEEHTIITDLPPLKLMHSFCAFKADGGPCRAIMKRFNFNITRQCEEFYGGCE 60
 DB 1 DSEDEEHTIITDLPPLKLMHSFCAFKADGGPCRAIMKRFNFNITRQCEEFYGGCE 60
 QY 61 GNONRFESLECKKMKCTRDANRIKTTLOQEKPDFCFLEEDPGICRGYITRYFYNNQTK 120
 DB 61 GNONRFESLECKKMKCTRDANRIKTTLOQEKPDFCFLEEDPGICRGYITRYFYNNQTK 120
 QY 121 QCERFYGGCLGNMNNFETLECKNICEGDPNGFQVDNYGT 161
 DB 121 QCERFYGGCLGNMNNFETLECKNICEGDPNGFQVDNYGT 161
 RESULT 4
 AAU02971
 ID AAU02971 standard; Protein; 225 AA.
 XX AC
 XX AC AAU02971;
 XX XX 12-SEP-2001 (first entry)
 XX DE
 XX DE Angiotensin converting enzyme (ACEV) splice variant protein #71.
 KW Angiotensin converting enzyme splice variant; ACEV; interleukin 6;
 KW granulocyte colony stimulating factor receptor; glucagon; hypertrophy;
 KW platelet-derived endothelial cell growth factor; cardiovascular disease;
 KW cellular tumour antigen P53; cyclin-dependent kinase inhibitor 1C;

KW vasoactive intestinal polypeptide receptor 2; arteriosclerosis; cancer;
 KW myocardial infarction; coronary arterial thrombosis; renal disease;
 KW diabetic nephropathy; muscular disease; immune disorder; sarcoidosis;
 KW multiple sclerosis; immune complex nephritis; deep vein thrombosis;
 KW nonaroidotic pulmonary granulomatous disease; endothelial abnormality;
 KW vascular disorder; asbestosis.
 XX OS Homo sapiens.
 XX XX WO200136632-A2.
 XX PN 25-MAY-2001.
 XX PD
 XX PF 17-NOV-2000; 2000WO-IL00766.
 XX XX
 XX PR 17-NOV-1999; 99IL-0132978.
 XX PR 10-DEC-1999; 99IL-0133455.
 XX XX
 XX PA (COMP-) COMPUGEN LTD.
 XX XX
 XX PI Levine Z, David A, Azar I, Khosravi R, Bernstein J;
 XX DR
 XX DR WPI; 2001-336004/35.
 XX DR N-PSDB; AAS06071.
 XX XX
 PT Novel alternative splicing variants e.g. variant of angiotensin
 PT converting enzyme (ACEV), useful in identifying candidate compounds
 PT capable of binding to the variant and to detect anti-variant antibodies
 XX PT
 XX PS
 XX PS Claim 4; Fig 71; 519pp; English.
 XX CC
 CC The sequence represents an angiotensin converting enzyme splice variant
 CC (ACEV) polypeptide. The polypeptides of the invention include variants of
 CC granulocyte colony stimulating factor receptor, glucagon, interleukin 6,
 CC platelet-derived endothelial cell growth factor, cyclin-dependent kinase
 CC inhibitor 1C, cellular tumour antigen P53, and vasoactive intestinal
 CC polypeptide receptor 2. The polypeptides and their associated nucleic
 CC acids are useful for identification of variant sequences and detection of
 CC candidate compounds capable of binding to the molecules. The sequences of
 CC the invention can be used in the treatment and diagnosis of various
 CC disorders including cardiovascular diseases such as arteriosclerosis,
 CC myocardial infarction and coronary arterial thrombosis, renal diseases
 CC such as diabetic nephropathy, muscular diseases such as hypertrophy,
 CC immune disorders such as immune complex nephritis, multiple sclerosis,
 CC cancer, sarcoidosis, nonaroidotic pulmonary granulomatous diseases such
 CC as asbestosis and vascular pathologies involving an endothelial
 CC abnormality such as deep vein thrombosis.
 XX CC
 XX SQ Sequence 225 AA;
 Query Match 99.1%; Score 897; DB 22; Length 225;
 Best Local Similarity 98.8%; Pred. No. 1.9e-82;
 Matches 159; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 DSEDEEHTIITDLPPLKLMHSFCAFKADGGPCRAIMKRFNFNITRQCEEFYGGCE 60
 DB 29 DSEDEEHTIITDLPPLKLMHSFCAFKADGGPCRAIMKRFNFNITRQCEEFYGGCE 88
 QY 61 GNONRFESLECKKMKCTRDANRIKTTLOQEKPDFCFLEEDPGICRGYITRYFYNNQTK 120
 DB 89 GNONRFESLECKKMKCTRDANRIKTTLOQEKPDFCFLEEDPGICRGYITRYFYNNQTK 148
 QY 121 QCERFYGGCLGNMNNFETLECKNICEGDPNGFQVDNYGT 161
 DB 149 QCERFYGGCLGNMNNFETLECKNICEGDPNGFQVDNYGT 189
 RESULT 5
 AAU03011
 ID AAU03011 standard; protein; 276 AA.
 XX AC
 XX AC AAU03011;

XX	30-JAN-1998	(first entry)
XX	Recombinant non-glycosylated TFPI.	
XX	Tissue factor pathway inhibitor; TFPI; human; tissue factor inhibitor;	
KW	Lipoprotein-associated coagulation inhibitor; coagulation inhibitor; TFI;	
KW	LACI; extrinsic pathway inhibitor; protein refolding; clot-inhibitor;	
KW	protein solubility modification; EPI.	
XX	Homo sapiens.	
OS		
XX		
XX		
PH	Key	
XX	Location/Qualifiers	
FT	Modified-site 2	
FT	/note= "phosphorylated to varying degrees, but does not	
FT	affect TFPI function"	
FT	Disulfide-bond 26..76	
FT	Disulfide-bond 35..59	
FT	Disulfide-bond 51..72	
FT	Disulfide-bond 97..147	
FT	Disulfide-bond 106..130	
FT	Disulfide-bond 122..143	
FT	Disulfide-bond 189..239	
FT	Disulfide-bond 198..222	
FT	Disulfide-bond 214..235	
XX		
XX	WO9640784-A2.	
PN		
XX		
PD	19-DEC-1996.	
XX		
PF	07-JUN-1996; 96WO-US09980.	
XX		
PR	07-JUN-1995; 95US-0477677.	
PR	07-JUN-1995; 95US-0473668.	
XX		
XX	(CHIR) CHIRON CORP.	
PA	(SEAR) SEARLE & CO G D.	
PA		
XX	Arve BH, Bild GS, Chen B, Dorin GJ, Gustafson ME;	
PI	Hallenbeck RF, Hora MS, Johnson GV, Johnson K, Madani H;	
PI	Pattison GL, Rana RK, Tsang M;	
XX		
XX	WPI; 1997-087056/08.	
DR		
XX		
XX	Aq. formulation of tissue factor pathway inhibitor - contains	
PT	charged polymer, e.g. dextran sulphate, to facilitate	
PT	solubilisation, formulation purification and refolding of protein	
PT	disclosure; Fig 4; 86pp; English.	
PS		
XX		
CC	This sequence represents the human tissue factor pathway inhibitor (TFPI)	
CC	protein. TFPI is also known as Lipoprotein-associated coagulation	
CC	inhibitor (LACI), extrinsic pathway inhibitor (EPI) and Tissue factor	
CC	inhibitor (TFI). The DNA encoding this sequence was altered to contain	
CC	the optimum codons for expression in E. coli, to allow for expression of	
CC	this protein in the bacteria. TFPI is used in the aqueous formulation of	
CC	the invention. The aqueous formulation also includes a charged polymer	
CC	(CP), preferably a sulphated polysaccharide (such as heparin or dextran	
CC	sulphate) or a polyphosphate, preferably immobilised on a solid support.	
CC	The CP is added to aid the correct refolding of TFPI. TFPI can also be	
CC	modified or refolded using the methods of the invention. One method is	
CC	for modifying the solubility of a protein, by adding an aqueous solution	
CC	of a CP to reduce inter- and intra-molecular interactions between the	
CC	charged domains of the protein. The second method is for refolding an	
CC	improperly folded or denatured protein (e.g. TFPI), and comprises adding	
CC	CP to a solution of the protein prior to allowing the protein to refold.	
CC	The methods are particularly useful for solubilising, formulating,	
CC	purifying and refolding proteins (especially TFPI) which have been	
CC	engineered by genetic recombination and produced in bacterial, yeast or	
CC	other cells in a form that has a non-native tertiary structure. TFPI is a	
CC	coagulation inhibitor which has clot-inhibiting properties.	
XX		
SQ	Sequence 276 AA;	
XX		

Query Match	99.1%;	Score 897;	DB 18;	Length 276;
Best Local Similarity	98.8%;	Pred. No. 2.4e-82;		
Matches 159;	Conservative 1;	Mismatches 1;	Indels 0;	Gaps 0;
QY	1	DSEDEEHTIIITDELPLPLKLMHSFAKADDDGPKRAIMKRRFFNFTTQCEEFYIGGCE	60	
Db	1	DSEDEEHTIIITDELPLPLKLMHSFAKADDDGPKRAIMKRRFFNFTTQCEEFYIGGCE	60	
QY	61	GNQNRFSLEECKKMCRTDNRANRIKTTTLOQEKPDPCFLEEDPGICRGVITRYFYNNQTK	120	
Db	61	GNQNRFSLEECKKMCRTDNRANRIKTTTLOQEKPDPCFLEEDPGICRGVITRYFYNNQTK	120	
QY	121	QCERFKYGGCLGNMNNPETLEEKNICEDGPNGFQVDNYGT	161	
Db	121	QCERFKYGGCLGNMNNPETLEEKNICEDGPNGFQVDNYGT	161	
RESULT 6				
AAW61535	ID	AAW61535 standard; protein; 276 AA.		
XX	AAW61535;			
XX	06-NOV-1998	(first entry)		
DT	06-NOV-1998	(first entry)		
XX	Human tissue factor pathway inhibitor (TFPI).			
DE	Human tissue factor pathway inhibitor; TFPI; TPPI-2; cell proliferation			
KW	angiogenesis-related disease; cancer; arthritis; macular degeneration			
KW	diabetic retinopathy.			
XX	Homo sapiens.			
XX	Key	Location/Qualifiers		
FH	Modified-site	2		
FT	/note=	"Potential phosphorylation site"		
FT	Domain	26..76		
FT	Disulfide-bond	26..76		
FT	Disulfide-bond	35..59		
FT	Disulfide-bond	51..72		
FT	Domain	97..147		
FT	/note=	"Kunitz-2 type protease inhibitor domain"		
FT	Disulfide-bond	97..147		
FT	Disulfide-bond	106..130		
FT	Modified-site	117		
FT	/note=	"N-glycosylated"		
FT	Disulfide-bond	122..143		
FT	Modified-site	167		
FT	/note=	"N-glycosylated"		
FT	Domain	189..239		
FT	/note=	"Kunitz-3 type protease inhibitor domain"		
FT	Disulfide-bond	189..239		
FT	Disulfide-bond	198..222		
FT	Disulfide-bond	214..235		
FT	Modified-site	228		
FT	/note=	"N-glycosylated"		
XX	WO9834634-A1.			
PN	13-AUG-1998.			
PD	13-AUG-1998.			
XX	06-FEB-1998;	98WO-0502699.		
PF	06-FEB-1998;	97US-0796850.		
XX	06-FEB-1997;	97US-0796850.		
PR	(ENTR-) ENTREMED INC.			
PA	Green SJ, Papathanassiou AE;			
XX	WPI; 1998-446947/38.			
PI				
XX				
DR				
XX				

PS Disclosure; Page 27-28; 56pp; English.

XX A Sali fragment encoding human tissue factor pathway inhibitor

CC (TFPI) is given in AA081396. It was used to construct expression

CC vectors allowing production of TFPI in *Saccharomyces cerevisiae* as

CC fusions to the yeast aspartic protease 3 (YAP3) signal peptide,

CC facilitating secretion of the recombinant TFPI.

XX

XX Sequence 304 AA;

Query Match 99.1%; Score 897; DB 16; Length 304;

Best Local Similarity 98.8%; Pred. No. 2.7e-82;

Matches 159; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 DSEDEEHTIITDTLPPLKLMHSCAFKADGPGCAIMKRRFFNFTQCEEFYGGCE 60

DB 29 DSEDEEHTIITDTLPPLKLMHSCAFKADGPGCAIMKRRFFNFTQCEEFYGGCE 88

QY 61 GNQNFESLECKKCTRDNRANRIKTTLQOEKDPDFCFLEEDPGICRGYITRYFYNNQTK 120

DB 89 GNQNFESLECKKCTRDNRANRIKTTLQOEKDPDFCFLEEDPGICRGYITRYFYNNQTK 148

QY 121 QCERFYGGCLGNMNFETLECKNICEGDPNGFQVDNYGT 161

DB 149 QCERFYGGCLGNMNFETLECKNICEGDPNGFQVDNYGT 189

RESULT 9

AAV49557

ID AAV49557 standard; Protein; 304 AA.

XX

AC AAV49557;

XX

XX 13-JAN-2000 (first entry)

XX Human lipoprotein associated coagulation inhibitor protein sequence.

XX

XX Human; coding sequence polymorphism; vascular pathology gene;

KW polymorphic site; phenotype correlation; forensic; paternity testing;

KW medicine; genetic analysis; vascular disease.

XX

OS Homo sapiens.

XX

XX WO950454-A2.

PN

PD 07-OCT-1999.

XX

PF 26-MAR-1999; 99WO-US06473.

XX

PR 01-APR-1998; 98US-0054272.

XX

PA (WHED) WHITEHEAD INST BIOMEDICAL RES.

XX

PI Lander ES, Daley GQ, Cargill M, Ireland JS, Rozen SG;

XX

DR WPI; 1999-620066/53.

XX

DR N-PSDB; AA232166.

XX

PT Determination of polymorphisms in genes, especially those identifying

PT predisposition to vascular disease

XX

PS Disclosure; Fig 10; 134pp; English.

XX

XX AA232159 to AA232194 represent reference alleles for specifically

CC claimed nucleic acid sequences from the present invention which comprise

CC polymorphic sites as given in a table in the specification, selected

CC from 92 single nucleotide polymorphisms in which the nucleotide at the

CC polymorphic site is different from a nucleotide at the same site in a

CC reference allele. The nucleic acids, and primers and probes, are used to

CC identify polymorphisms, which may predispose an individual to disease,

CC especially a vascular disease. They can also be used in phenotype

CC correlations, forensics, paternity testing, medicine or genetic

CC analysis. AAV49550 to AAV49573 represent the proteins which correspond

CC

CC to some of the reference alleles.

XX

XX Sequence 304 AA;

Query Match 99.1%; Score 897; DB 20; Length 304;

Best Local Similarity 98.8%; Pred. No. 2.7e-82;

Matches 159; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 DSEDEEHTIITDTLPPLKLMHSCAFKADGPGCAIMKRRFFNFTQCEEFYGGCE 60

DB 29 DSEDEEHTIITDTLPPLKLMHSCAFKADGPGCAIMKRRFFNFTQCEEFYGGCE 88

QY 61 GNQNFESLECKKCTRDNRANRIKTTLQOEKDPDFCFLEEDPGICRGYITRYFYNNQTK 120

DB 89 GNQNFESLECKKCTRDNRANRIKTTLQOEKDPDFCFLEEDPGICRGYITRYFYNNQTK 148

QY 121 QCERFYGGCLGNMNFETLECKNICEGDPNGFQVDNYGT 161

DB 149 QCERFYGGCLGNMNFETLECKNICEGDPNGFQVDNYGT 189

RESULT 10

AAV92011

ID AAV92011 standard; Protein; 352 AA.

XX

AC AAV92011;

XX

XX 08-MAY-1996 (first entry)

XX Ubiquitin-TFPI fusion protein.

XX

XX Tissue factor pathway inhibitor; TFPI; TFPI-2; Factor-VIIa;

KW tissue factor; Factor-Xa; binding protein; tissue factor inhibitor;

KW lipoprotein associated coagulation inhibitor;

KW extrinsic pathway inhibitor; sepsis; septic shock;

KW *Saccharomyces cerevisiae*.

XX

OS Chimeric synthetic;

OS Chimeric Homo sapiens.

XX

XX Key Location/Qualifiers

FT Protein 1..73

FT /note= "amino acids 1-73 comprise the ubiquitin

FT fusion partner"

FT Peptide 74..76

FT /note= "amino acids 74-76 comprise a linker

FT sequence"

FT Protein 77..352

FT /note= "amino acids 77-352 comprise mature TFPI"

XX

XX WO9604377-A1.

PN

XX 15-FEB-1996.

PD

XX 25-JUL-1995; 95WO-US09377.

PF

XX 05-AUG-1994; 94US-0286530.

XX

XX (CHIR) CHIRON CORP.

PA

XX Creasey AA, Innis MA;

PI

XX WPI; 1996-129393/13.

DR

DR N-PSDB; AAT16056.

XX

XX Production of tissue factor pathway inhibitor in yeast cells - with

PT isolation from the insoluble cell fraction, used to treat or prevent

PT sepsis or septic shock

XX

XX Disclosure; Page 20-22; 36pp; English.

PS

XX A ubiquitin/tissue factor pathway inhibitor (TFPI) fusion protein

CC (AAV92011) is produced within *Saccharomyces cerevisiae* host cells

CC

CC following transformation with vector plasmid pLACI 4.1, which carries a ubiquitin/TFPI gene fusion (AAT16056). The fusion protein is cleaved within the yeast cell, so that mature TFPI having an authentic N-terminal sequence (see AAR92013) is obtained. Max. prodn. of TFPI from yeast transformants reached 1 mg/ml. The TFPI is used to treat or prevent sepsis and septic shock.

XX
SQ Sequence 352 AA;

Query Match 99.1%; Score 897; DB 17; Length 352;
Best Local Similarity 98.8%; Pred. No. 3.3e-82;
Matches 159; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 DSEDEEHTIITDTLPPLKLMHSFCFAKADGDCRAIMKRRFFNFTQCEEFYGGCE 60
DB 77 DSEDEEHTIITDTLPPLKLMHSFCFAKADGDCRAIMKRRFFNFTQCEEFYGGCE 136
QY 61 GGNRFESLECKKMCCTRONANRIKTTLQOEKPDFCFLEEDPGICRGYITRYFYNNQTK 120
DB 137 GGNRFESLECKKMCCTRONANRIKTTLQOEKPDFCFLEEDPGICRGYITRYFYNNQTK 196
QY 121 QCFERFYGGCLGNMNFETLECKNICEGDPNGFQVDNYGT 161
DB 197 QCFERFYGGCLGNMNFETLECKNICEGDPNGFQVDNYGT 237

RESULT 11
AAR23800
ID AAR23800 standard; Protein; 160 AA.
XX
AC AAR23800;
XX
DT 03-NOV-1992 (first entry)
XX
DE LACI fragment 1 - 160.
XX
KW Lipoprotein-associated coagulation inhibitor; kuintz.
XX
OS Homo sapiens.
XX
PN US5106833-A.
XX
PD 21-APR-1992.
XX
PF 26-JAN-1989; 89US-0301779.
XX
PR 23-JUL-1987; 87US-0077366.
XX
PR 23-NOV-1987; 87US-0123753.
XX
PR 26-JAN-1989; 89US-0301779.
XX
PA (UNIW) UNIV WASHINGTON.
XX
PI Broze GJ, Girard TJ;
XX
DR WPI; 1992-159405/19.
XX
PT Peptide fragments of lipoprotein-associated coagulation inhibitor
PT - used for inhibiting Factor Xa prodn. or inhibiting Factor VIIA
PT tissue factor complex formation
XX
PS Disclosure; Fig 1; 9pp; English.
XX
CC The sequence given is a peptide fragment of lipoprotein-associated
CC coagulation inhibitor (LACI) having the sequence of residues 1
CC to 160 of the 276 residue mature LACI protein. This peptide fragment
CC corresponds to Kunitz domains of LACI and can be used for
CC inhibiting Factor VIIA/tissue factor enzymatic complex formation.
CC Factor VIIa is the activated form of factor VII.
XX
SQ Sequence 160 AA;

Query Match 98.6%; Score 892; DB 13; Length 160;
Best Local Similarity 98.8%; Pred. No. 4.1e-82;

Matches 158; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 DSEDEEHTIITDTLPPLKLMHSFCFAKADGDCRAIMKRRFFNFTQCEEFYGGCE 60
DB 1 DSEDEEHTIITDTLPPLKLMHSFCFAKADGDCRAIMKRRFFNFTQCEEFYGGCE 60
QY 61 GGNRFESLECKKMCCTRONANRIKTTLQOEKPDFCFLEEDPGICRGYITRYFYNNQTK 120
DB 61 GGNRFESLECKKMCCTRONANRIKTTLQOEKPDFCFLEEDPGICRGYITRYFYNNQTK 120
QY 121 QCFERFYGGCLGNMNFETLECKNICEGDPNGFQVDNYGT 160
DB 121 QCFERFYGGCLGNMNFETLECKNICEGDPNGFQVDNYGT 160

RESULT 12
AAP92002
ID AAP92002 standard; protein; 304 AA.
XX
AC AAP92002;
XX
DT 09-FEB-1990 (first entry)
XX
DE Human tissue factor inhibitor (TFI).
XX
KW Human tissue factor inhibitor; TFI; human placenta lambda-P9 clone; basic
KW protease inhibitor gene superfamily
XX
OS Homo sapiens.
XX
FH Key
FH Region 1..28
FT /label= Signal_region
FT /note= "A-T rich"
FT Cleavage-site 28..29
FT /note= "Possible site for signal peptidase"
FT Protein 29..304
FT Region 145..147
FT /note= "Potential N-linked glycosylation site"
FT Region 195..197
FT /note= "Potential N-linked glycosylation site"
FT Region 256..258
FT /note= "Potential N-linked glycosylation site"
FT Region 31..53
FT /note= "See comments below"
FT Region 82..88
FT /note= "See comments below"
FT Region 153..16
FT /note= "See comments below"
XX
PN EP318451-A.
XX
PD 31-MAY-1989.
XX
PF 22-JUL-1988; 88EP-0870127.
XX
PR 23-NOV-1987; 87US-0123753.
XX
PA (MONS) MONSANTO CO (UNIW).
XX
PI Broze GJ, Kretzmer KK, Wun TC;
XX
DR WPI; 1989-159483/22.
XX
DR N-PSDB; AAN90108.
XX
PT DNA encoding human tissue factor inhibitor
PT - used in study of coagulation cascade for agents
PT which inhibit factor Xa and Factor VIIA-TF
XX
XX Claim 3; Figure 3; 14pp; English.
XX
CC Amino acid sequence of tissue factor inhibitor (TFI) cDNA isolate from
CC lambda-P9 clone of human placenta cDNA library. Domains discerned

CC include: highly negatively charged N-terminal; highly positively charged
CC carboxy-terminal; intervening portion consisting of 3 homologous domains
CC with sequences typical of Kunitz-type enzyme inhibitors. Based on
CC homology study, it appears to be a member of the basic protease
CC inhibitor gene superfamily. Sequences in misc. regions in feature table
CC above have been independently confirmed by amino acid sequence analysis.

XX SQ Sequence 304 AA;

Query Match 98.5%; Score 891; DB 10; Length 304;
Best Local Similarity 98.1%; Pred. No. 1.1e-81;
Matches 158; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 DSEDEEHTIITDTLPPLKLMHSFCFAKADGPGCAIMKRRFFNFTQCEEFYGGCE 60
DB 29 DSEDEEHTIITDTLPPLKLMHSFCFAKADGPGCAIMKRRFFNFTQCEEFYGGCE 88
QY 61 GNQRFESLECKKMKCTRDNANRIKTTLOQEKDPDFCFLEEDPGICRGYITRYFYNNQTK 120
DB 89 GNQRFESLECKKMKCTRDNANRIKTTLOQEKDPDFCFLEEDPGICRGYITRYFYNNQTK 148
QY 121 QCERFKYGGCLGNMNNFETLECKNICEGDPNGFOVDNYGT 161
DB 149 QCERFKYGGCLGNMNNFETLECKNICEGDPNGFOVDNYGT 189

RESULT 13

AAR11169
ID AAR11169 standard; Protein; 256 AA.

AC AAR11169;

DT 21-MAY-1991 (first entry)

DE Ser(Asp1-Thr255)-EPI analogue.

KW Extrinsic pathway inhibitor protein; tissue plasminogen activator;
KW kunitz domain; coagulation.

OS Synthetic.

PN WO9102753-A.

PD 07-MAR-1991.

PF 17-AUG-1990; 90WO-DK00212.

PR 18-AUG-1989; 89DK-0004080.

PA (NOVO) NOVO NORDISK A/S.

PI Rasmussen J, Nordfang O;

DR WPI; 1991-087248/12.

PT Extrinsic pathway inhibitor protein analogue - useful as
PT anticoagulant and anti-cancer agent due to low or no
PT heparin-binding capacity

PS Claim 5; Page 20; 39pp; English.

CC Preferred EPI analogues of the invention comprise the first two
CC kunitz domains of native EPI and have one or more amino acids
CC deleted in the region Glu148 to the C-terminal Met276, especially
CC in the region from Arg246 to Lys275. This analogue is an example of
CC such a peptide; it corresponds to amino acids 1 to 255 of native EPI
CC with an additional N-terminal Ser residue. The analogue can be used
CC in a therapeutic composition to treat patients having coagulation
CC disorders or cancer. The analogue has a longer half-life than
CC full-length EPI.

CC See also AAQ10992-4 and AAR11170-R11172.

XX Sequence 256 AA;

Query Match 98.3%; Score 890; DB 12; Length 256;
Best Local Similarity 98.1%; Pred. No. 1.1e-81;
Matches 158; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 DSEDEEHTIITDTLPPLKLMHSFCFAKADGPGCAIMKRRFFNFTQCEEFYGGCE 60
DB 2 DSEDEEHTIITDTLPPLKLMHSFCFAKADGPGCAIMKRRFFNFTQCEEFYGGCE 61
QY 61 GNQRFESLECKKMKCTRDNANRIKTTLOQEKDPDFCFLEEDPGICRGYITRYFYNNQTK 120
DB 62 GNQRFESLECKKMKCTRDNANRIKTTLOQEKDPDFCFLEEDPGICRGYITRYFYNNQTK 121
QY 121 QCERFKYGGCLGNMNNFETLECKNICEGDPNGFOVDNYGT 161
DB 122 QCERFKYGGCLGNMNNFETLECKNICEGDPNGFOVDNYGT 162

RESULT 14

AAR11171
ID AAR11171 standard; Protein; 261 AA.

AC AAR11171;

DT 21-MAY-1991 (first entry)

DE Ser(Asp1-Glu245)-(Glu262-Met276)-EPI analogue.

KW Extrinsic pathway inhibitor protein; tissue plasminogen activator;
KW kunitz domain; coagulation.

OS Synthetic.

PN WO9102753-A.

PD 07-MAR-1991.

PF 17-AUG-1990; 90WO-DK00212.

PR 18-AUG-1989; 89DK-0004080.

PA (NOVO) NOVO NORDISK A/S.

PI Rasmussen J, Nordfang O;

DR WPI; 1991-087248/12.

PT Extrinsic pathway inhibitor protein analogue - useful as
PT anticoagulant and anti-cancer agent due to low or no
PT heparin-binding capacity

PS Claim 5; Page 20; 39pp; English.

CC Preferred EPI analogues of the invention comprise the first two
CC kunitz domains of native EPI and have one or more amino acids
CC deleted in the region Glu148 to the C-terminal Met276, especially
CC in the region from Arg246 to Lys275. This analogue is an example of
CC such a peptide; it has amino acids 246 to 261 of native EPI deleted
CC with an optional N-terminal Ser residue. The analogue can be used
CC in a therapeutic composition to treat patients having coagulation
CC disorders or cancer. The analogue has a longer half-life than
CC full-length EPI.

CC See also AAQ10992-4 and AAR11169-R11170, AAR11172.

XX Sequence 261 AA;

Query Match 98.3%; Score 890; DB 12; Length 261;
Best Local Similarity 98.1%; Pred. No. 1.2e-81;
Matches 158; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 DSEDEEHTIITDTLPPLKLMHSFCFAKADGPGCAIMKRRFFNFTQCEEFYGGCE 60
DB 2 DSEDEEHTIITDTLPPLKLMHSFCFAKADGPGCAIMKRRFFNFTQCEEFYGGCE 61

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model
Run on: December 2, 2002, 10:02:30 ; Search time 26.4328 Seconds
(without alignments)
179.212 Million cell updates/sec

Title: US-09-741-106-19
Perfect score: 905
Sequence: 1 DSEDEHTITDTELPPLK.....ECKNICDGPNGFQVDNYGT 161

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/1/iaa/5A.COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B.COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A.COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B.COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCUS.COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	905	100.0	161	1	US-08-437-841-19
2	905	100.0	161	1	US-08-286-521-19
3	905	100.0	161	1	US-08-436-175-19
4	905	100.0	161	4	US-08-943-682-19
5	905	100.0	161	5	PCT-US95-09464-19
6	900	99.4	276	1	US-08-437-841-9
7	900	99.4	276	1	US-08-286-521-9
8	900	99.4	276	1	US-08-436-175-9
9	900	99.4	276	3	US-08-854-764-3
10	900	99.4	276	4	US-08-943-682-9
11	900	99.4	276	5	PCT-US95-09377-3
12	900	99.4	276	5	PCT-US95-09464-9
13	897	99.1	276	2	US-08-796-850-1
14	897	99.1	277	1	US-07-844-297-1
15	897	99.1	304	1	US-08-026-145-2
16	897	99.1	304	1	US-08-446-646-9
17	897	99.1	304	1	US-08-676-125A-18
18	897	99.1	304	2	US-09-136-012A-18
19	897	99.1	304	3	US-08-676-124-1
20	897	99.1	304	3	US-08-208-264A-25
21	897	99.1	304	3	US-09-414-878-1
22	897	99.1	304	3	US-09-240-136-1
23	897	99.1	304	4	US-09-421-097-25
24	897	99.1	304	4	US-09-638-770A-1
25	897	99.1	304	6	5466783-2
26	897	99.1	352	3	US-08-854-764-2
27	897	99.1	352	5	PCT-US95-09377-2

28	891	98.5	189	1	US-07-828-920A-7	Sequence 7, Appli
29	891	98.5	304	4	US-09-054-782-2	Sequence 2, Appli
30	890	98.3	276	1	US-07-828-920A-1	Sequence 1, Appli
31	825	91.2	183	1	US-07-828-920A-5	Sequence 5, Appli
32	787	87.0	213	6	5466783-25	Patent No. 5466783
33	425	47.0	110	1	US-07-828-920A-3	Sequence 3, Appli
34	347.5	38.4	111	1	US-08-321-658B-12	Sequence 12, Appli
35	334	36.9	58	1	US-08-321-658B-8	Sequence 8, Appli
36	334	36.9	58	1	US-08-384-489-14	Sequence 14, Appli
37	334	36.9	58	1	US-08-358-160-76	Sequence 76, Appli
38	334	36.9	58	1	US-08-463-155A-36	Sequence 36, Appli
39	334	36.9	58	1	US-08-463-432B-36	Sequence 36, Appli
40	334	36.9	58	1	US-08-676-125A-33	Sequence 33, Appli
41	334	36.9	58	1	US-08-206-310A-36	Sequence 36, Appli
42	334	36.9	58	2	US-08-398-010A-36	Sequence 36, Appli
43	334	36.9	58	2	US-08-398-628A-36	Sequence 36, Appli
44	334	36.9	58	2	US-08-399-115A-36	Sequence 36, Appli
45	334	36.9	58	2	US-09-136-012A-33	Sequence 33, Appli

ALIGNMENTS

RESULT 1
US-08-437-841-19
; Sequence 19, Application US/08437841
; Patent No. 5563123
; GENERAL INFORMATION:
; APPLICANT: Innis, Michael
; TITLE OF INVENTION: Chimeric Proteins
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton St.
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/437,841
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/286,521
; FILING DATE: 05-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Saveriede, Paul B.
; REGISTRATION NUMBER: 36,914
; REFERENCE/DOCKET NUMBER: 0990.001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 510-601-2585
; TELEFAX: 510-655-3542
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 161 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-437-841-19

Query Match 100.0%; Score 905; DB 1; Length 161;
Best Local Similarity 100.0%; Pred. No. 1.1e-92;
Matches 161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DSEDEHTITDTELPPLKMHSCAFKADGPGCAIMKRRFFFTTQCEFFIYCGGE 60
|||||

Db 1 DSEDEEHTITDTLPLKLMHSFCAFKADDDGPCRAIMKRRFFNITRQCEEFYGGCE 60
Qy 61 GNQRFESLECKKMCRTDNANRIKTTLQOEKPDFCFLEEDPGICRGYITRYFYNQTK 120
Db 61 GNQRFESLECKKMCRTDNANRIKTTLQOEKPDFCFLEEDPGICRGYITRYFYNQTK 120
Qy 121 QCFERFYGGCLGNMNNFETLECKNCEDEGPNFGQVDNYGT 161
Db 121 QCFERFYGGCLGNMNNFETLECKNCEDEGPNFGQVDNYGT 161

RESULT 2

US-08-286-521-19
Sequence 19, Application US/08286521
Patent No. 5589359
GENERAL INFORMATION:
APPLICANT: Innis, Michael
TITLE OF INVENTION: Chimeric Proteins
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Chiron Corporation
STREET: 4560 Horton St.
CITY: Emeryville
STATE: CA
COUNTRY: USA
ZIP: 94608
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30B
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/286,521
FILING DATE: 05-AUG-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Savereide, Paul B.
REGISTRATION NUMBER: 36,914
REFERENCE/DOCKET NUMBER: 0990.001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 510-601-2585
TELEFAX: 510-655-3542
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 161 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-286-521-19

Query Match 100.0%; Score 905; DB 1; Length 161;
Best Local Similarity 100.0%; Pred. No. 1.1e-92;
Matches 161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 DSEDEEHTITDTLPLKLMHSFCAFKADDDGPCRAIMKRRFFNITRQCEEFYGGCE 60
Db 1 DSEDEEHTITDTLPLKLMHSFCAFKADDDGPCRAIMKRRFFNITRQCEEFYGGCE 60
Qy 61 GNQRFESLECKKMCRTDNANRIKTTLQOEKPDFCFLEEDPGICRGYITRYFYNQTK 120
Db 61 GNQRFESLECKKMCRTDNANRIKTTLQOEKPDFCFLEEDPGICRGYITRYFYNQTK 120
Qy 121 QCFERFYGGCLGNMNNFETLECKNCEDEGPNFGQVDNYGT 161
Db 121 QCFERFYGGCLGNMNNFETLECKNCEDEGPNFGQVDNYGT 161

RESULT 3

US-08-436-175-19
Sequence 19, Application US/08436175
Patent No. 5696088

GENERAL INFORMATION:
APPLICANT: Innis, Michael
TITLE OF INVENTION: Chimeric Proteins
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Chiron Corporation
STREET: 4560 Horton St.
CITY: Emeryville
STATE: CA
COUNTRY: USA
ZIP: 94608
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30B
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/436,175
FILING DATE: 09-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,521
FILING DATE: 05-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Savereide, Paul B.
REGISTRATION NUMBER: 36,914
REFERENCE/DOCKET NUMBER: 0990.001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 510-601-2585
TELEFAX: 510-655-3542
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 161 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-436-175-19

Query Match 100.0%; Score 905; DB 1; Length 161;
Best Local Similarity 100.0%; Pred. No. 1.1e-92;
Matches 161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 DSEDEEHTITDTLPLKLMHSFCAFKADDDGPCRAIMKRRFFNITRQCEEFYGGCE 60
Db 1 DSEDEEHTITDTLPLKLMHSFCAFKADDDGPCRAIMKRRFFNITRQCEEFYGGCE 60
Qy 61 GNQRFESLECKKMCRTDNANRIKTTLQOEKPDFCFLEEDPGICRGYITRYFYNQTK 120
Db 61 GNQRFESLECKKMCRTDNANRIKTTLQOEKPDFCFLEEDPGICRGYITRYFYNQTK 120
Qy 121 QCFERFYGGCLGNMNNFETLECKNCEDEGPNFGQVDNYGT 161
Db 121 QCFERFYGGCLGNMNNFETLECKNCEDEGPNFGQVDNYGT 161

RESULT 4

US-08-943-682-19
Sequence 19, Application US/08943682
Patent No. 6174721
GENERAL INFORMATION:
APPLICANT: Innis, Michael
TITLE OF INVENTION: Chimeric Proteins
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Chiron Corporation
STREET: 4560 Horton St.
CITY: Emeryville
STATE: CA
COUNTRY: USA
ZIP: 94608


```
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/943,682
; FILING DATE: 03-OCT-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/438,184
; FILING DATE: 09-MAY-1995
; APPLICATION NUMBER: US 08/286,521
; FILING DATE: 05-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Savereide, Paul B.
; REGISTRATION NUMBER: 36,914
; REFERENCE/DOCKET NUMBER: 0990.001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 510-601-2585
; TELEFAX: 510-655-3542
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 161 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-943-682-19

Query Match 100.0%; Score 905; DB 4; Length 161;
Best Local Similarity 100.0%; Pred. No. 1.1e-92;
Matches 161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DSEDEEHTIITDTPPLKLMHSFCAKADGPGCAIMKRRFFNFTROCEEFIYGGCE 60
Db 1 DSEDEEHTIITDTPPLKLMHSFCAKADGPGCAIMKRRFFNFTROCEEFIYGGCE 60
QY 61 GNQRFESLECKKMCCTRDNRANRIKTTLQOEKPDFCFLEEDPGICRGYITRYFYNQOTK 120
Db 61 GNQRFESLECKKMCCTRDNRANRIKTTLQOEKPDFCFLEEDPGICRGYITRYFYNQOTK 120
QY 121 QCERFYGGCLGNMNFETLECKNICEDGPNQFQVDNYGT 161
Db 121 QCERFYGGCLGNMNFETLECKNICEDGPNQFQVDNYGT 161

RESULT 5
PCT-US95-09464-19
; Sequence 19, Application PC/TUS9509464
; GENERAL INFORMATION:
; APPLICANT: CHIRON CORPORATION
; TITLE OF INVENTION: Chimeric Proteins
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton St.
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/09464
; FILING DATE: 25-JULY-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Savereide, Paul B.
; REGISTRATION NUMBER: 36,914
```

```
;
; REFERENCE/DOCKET NUMBER: 0990.100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 510-601-2585
; TELEFAX: 510-655-3542
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 161 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; PCT-US95-09464-19

Query Match 100.0%; Score 905; DB 5; Length 161;
Best Local Similarity 100.0%; Pred. No. 1.1e-92;
Matches 161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DSEDEEHTIITDTPPLKLMHSFCAKADGPGCAIMKRRFFNFTROCEEFIYGGCE 60
Db 1 DSEDEEHTIITDTPPLKLMHSFCAKADGPGCAIMKRRFFNFTROCEEFIYGGCE 60
QY 61 GNQRFESLECKKMCCTRDNRANRIKTTLQOEKPDFCFLEEDPGICRGYITRYFYNQOTK 120
Db 61 GNQRFESLECKKMCCTRDNRANRIKTTLQOEKPDFCFLEEDPGICRGYITRYFYNQOTK 120
QY 121 QCERFYGGCLGNMNFETLECKNICEDGPNQFQVDNYGT 161
Db 121 QCERFYGGCLGNMNFETLECKNICEDGPNQFQVDNYGT 161

RESULT 6
US-08-437-841-9
; Sequence 9, Application US/08437841
; Patent No. 5563123
; GENERAL INFORMATION:
; APPLICANT: Innis, Michael
; APPLICANT: Creasey, Abba
; TITLE OF INVENTION: Chimeric Proteins
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton St.
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/437,841
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/286,521
; FILING DATE: 05-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Savereide, Paul B.
; REGISTRATION NUMBER: 36,914
; REFERENCE/DOCKET NUMBER: 0990.001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 510-601-2585
; TELEFAX: 510-655-3542
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 276 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-437-841-9
```



```
;
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton St.
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/854,764
; FILING DATE: 12-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/286,530
; FILING DATE: 05-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Savereide, Paul B.
; REGISTRATION NUMBER: 36,914
; REFERENCE/DOCKET NUMBER: 0991.001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 510-601-2585
; TELEFAX: 510-655-3542
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 276 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-854-764-3

Query Match 99.4%; Score 900; DB 3; Length 276;
Best Local Similarity 99.4%; Pred. No. 7.7e-92;
Matches 160; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DSEDEEHTITDTELPLPLKLMHSFCADDDGPCRAIMKRRFFNFTTQCEEFYGGCE 60
Db 1 DSEDEEHTITDTELPLPLKLMHSFCADDDGPCRAIMKRRFFNFTTQCEEFYGGCE 60

QY 61 GQNRFSLECKKMCRTDNANRIKTTLQEKPDGFCFLEEDPGICRGYITRYFYNNQTK 120
Db 61 GQNRFSLECKKMCRTDNANRIKTTLQEKPDGFCFLEEDPGICRGYITRYFYNNQTK 120

QY 121 QCERFYGGCLGNMNNFETLECKNICEGDPNGFQVDNYGT 161
Db 121 QCERFYGGCLGNMNNFETLECKNICEGDPNGFQVDNYGT 161

RESULT 10
US-08-682-9
; Sequence 9, Application US/08943682
; Patent No. 6174721
; GENERAL INFORMATION:
; APPLICANT: Innis, Michael
; TITLE OF INVENTION: Chimeric Proteins
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton St.
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

;
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton St.
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/438,184
; FILING DATE: 09-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/286,521
; FILING DATE: 05-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Savereide, Paul B.
; REGISTRATION NUMBER: 36,914
; REFERENCE/DOCKET NUMBER: 0990.001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 510-601-2585
; TELEFAX: 510-655-3542
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 276 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-943-682-9

Query Match 99.4%; Score 900; DB 4; Length 276;
Best Local Similarity 99.4%; Pred. No. 7.7e-92;
Matches 160; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DSEDEEHTITDTELPLPLKLMHSFCADDDGPCRAIMKRRFFNFTTQCEEFYGGCE 60
Db 1 DSEDEEHTITDTELPLPLKLMHSFCADDDGPCRAIMKRRFFNFTTQCEEFYGGCE 60

QY 61 GQNRFSLECKKMCRTDNANRIKTTLQEKPDGFCFLEEDPGICRGYITRYFYNNQTK 120
Db 61 GQNRFSLECKKMCRTDNANRIKTTLQEKPDGFCFLEEDPGICRGYITRYFYNNQTK 120

QY 121 QCERFYGGCLGNMNNFETLECKNICEGDPNGFQVDNYGT 161
Db 121 QCERFYGGCLGNMNNFETLECKNICEGDPNGFQVDNYGT 161

RESULT 11
PCT-US95-09377-3
; Sequence 3, Application PC/TUS9509377
; GENERAL INFORMATION:
; APPLICANT: CHIRON CORPORATION
; TITLE OF INVENTION: Production of Tissue Factor Pathway
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton St.
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/09377
; FILING DATE: 25-JULY-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Savereide, Paul B.
; REGISTRATION NUMBER: 36,914
; REFERENCE/DOCKET NUMBER: 0991.100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 510-601-2585
```

TELEFAX: 510-655-3542
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 276 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US95-09377-3

Query Match 99.4%; Score 900; DB 5; Length 276;
Best Local Similarity 99.4%; Pred. No. 7.7e-92;
Matches 160; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DSEDEEHTIITDTELPLKLMHSFCFAKADGDCRAIMKRRFFNIFTRQCEEFYGGCE 60
Db 1 DSEDEEHTIITDTELPLKLMHSFCFAKADGDCRAIMKRRFFNIFTRQCEEFYGGCE 60
QY 61 GNONRFESLECKKMCRTDNANRIKTTLOQEKPDFCFLEEDPGICRGYITRYFYNNQTK 120
Db 61 GNONRFESLECKKMCRTDNANRIKTTLOQEKPDFCFLEEDPGICRGYITRYFYNNQTK 120
QY 121 Q CERFKYGGCLGNMNFETLECKNICEDGPNQFQVDNYGT 161
Db 121 Q CERFKYGGCLGNMNFETLECKNICEDGPNQFQVDNYGT 161

RESULT 12

PCT-US95-09464-9
Sequence 9, Application PC/TUS9509464
GENERAL INFORMATION:
APPLICANT: CHIRON CORPORATION
TITLE OF INVENTION: Chimeric Proteins
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Chiron Corporation
STREET: 4560 Horton St.
CITY: Emeryville
STATE: CA
COUNTRY: USA
ZIP: 94608
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30B
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/09464
FILING DATE: 25-JULY-1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Savereide, Paul B.
REGISTRATION NUMBER: 36,914
REFERENCE/DOCKET NUMBER: 0990.100
TELECOMMUNICATION INFORMATION:
TELEPHONE: 510-601-2585
TELEFAX: 510-655-3542
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 276 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US95-09464-9

Query Match 99.4%; Score 900; DB 5; Length 276;
Best Local Similarity 99.4%; Pred. No. 7.7e-92;
Matches 160; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DSEDEEHTIITDTELPLKLMHSFCFAKADGDCRAIMKRRFFNIFTRQCEEFYGGCE 60
Db 1 DSEDEEHTIITDTELPLKLMHSFCFAKADGDCRAIMKRRFFNIFTRQCEEFYGGCE 60

QY 61 GNONRFESLECKKMCRTDNANRIKTTLOQEKPDFCFLEEDPGICRGYITRYFYNNQTK 120
Db 61 GNONRFESLECKKMCRTDNANRIKTTLOQEKPDFCFLEEDPGICRGYITRYFYNNQTK 120
QY 121 Q CERFKYGGCLGNMNFETLECKNICEDGPNQFQVDNYGT 161
Db 121 Q CERFKYGGCLGNMNFETLECKNICEDGPNQFQVDNYGT 161

RESULT 13

US-08-796-850-1
Sequence 1, Application US/08796850
Patent No. 5981471
GENERAL INFORMATION:
APPLICANT: Papathanassiou, Adonia E
APPLICANT: Green, Shawn J.
TITLE OF INVENTION: Compositions and Methods for Inhibiting
TITLE OF INVENTION: Cellular Proliferation
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Jones & Askew
STREET: 191 Peachtree Street, 37th Floor
CITY: Atlanta
STATE: Georgia
COUNTRY: U.S.A.
ZIP: 30303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/796,850
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Greene, Jamie L.
REGISTRATION NUMBER: 32,467
REFERENCE/DOCKET NUMBER: 05213-0290
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404) 818-3700
TELEFAX: (404) 818-3799
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 276 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: Active-site
LOCATION: 2..3
OTHER INFORMATION: /note= "Site of partial
OTHER INFORMATION: phosphorylation"
FEATURE:
NAME/KEY: Active-site
LOCATION: 117..118
OTHER INFORMATION:
OTHER INFORMATION: glycosylation"
FEATURE:
NAME/KEY: Active-site
LOCATION: 167..168
OTHER INFORMATION: /note= "Potential site for N-linked
OTHER INFORMATION: glycosylation"
FEATURE:
NAME/KEY: Active-site
LOCATION: 228..229

OTHER INFORMATION: /note= "Potential site for N-linked glycosylation"

FEATURE:

NAME/KEY: Domain

LOCATION: 26..76

OTHER INFORMATION: /label= Kunitz-1

FEATURE:

NAME/KEY: Domain

LOCATION: 97..147

OTHER INFORMATION: /label= Kunitz-2

FEATURE:

NAME/KEY: Domain

LOCATION: 189..239

OTHER INFORMATION: /label= Kunitz-3

US-08-796-850-1

Query Match 99.1%; Score 897; DB 2; Length 276;

Best Local Similarity 98.8%; Pred. No. 1.7e-91;

Matches 159; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 DSEDEHTIITDTELPPLKLMHSFCFAKADGPGCAIMKRFFNFTRQCEEFYGGCE 60

Db 1 DSEDEHTIITDTELPPLKLMHSFCFAKADGPGCAIMKRFFNFTRQCEEFYGGCE 60

QY 61 GNQRFESLECKKMCRTDNANRIKTTLQOEKPDGFCLEEDPGICRGYITRYFYNNQTK 120

Db 61 GNQRFESLECKKMCRTDNANRIKTTLQOEKPDGFCLEEDPGICRGYITRYFYNNQTK 120

QY 121 QCFERFYGGCLGNMNFETLECKNICEDGPNQFQVDNYGT 161

Db 121 QCFERFYGGCLGNMNFETLECKNICEDGPNQFQVDNYGT 161

RESULT 14

US-07-844-297-1

Sequence 1, Application US/07844297

Patent No. 5212091

GENERAL INFORMATION:

APPLICANT: Diaz-Collier, Judy A.

APPLICANT: Gustafson, Mark E.

APPLICANT: Wun, Tze-Chen

TITLE OF INVENTION: Method of Producing Tissue Factor

TITLE OF INVENTION: Pathway Inhibitor

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

ADDRESSEE: Scott J. Meyer, Monsanto Co., A3SG

STREET: 800 N. Lindbergh Blvd.

CITY: St. Louis

STATE: Missouri

COUNTRY: U.S.A.

ZIP: 63167

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/844,297

FILING DATE: 19920302

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Meyer, Scott J.

REGISTRATION NUMBER: 25,275

REFERENCE/DOCKET NUMBER: 07-21(819)A

TELEPHONE: (314)694-3117

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 277 amino acids

TYPE: AMINO ACID

TOPOLOGY: linear

MOLECULE TYPE: protein

US-07-844-297-1

Query Match 99.1%; Score 897; DB 1; Length 277;

Best Local Similarity 98.8%; Pred. No. 1.7e-91;

Matches 159; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 DSEDEHTIITDTELPPLKLMHSFCFAKADGPGCAIMKRFFNFTRQCEEFYGGCE 60

Db 2 DSEDEHTIITDTELPPLKLMHSFCFAKADGPGCAIMKRFFNFTRQCEEFYGGCE 61

QY 61 GNQRFESLECKKMCRTDNANRIKTTLQOEKPDGFCLEEDPGICRGYITRYFYNNQTK 120

Db 62 GNQRFESLECKKMCRTDNANRIKTTLQOEKPDGFCLEEDPGICRGYITRYFYNNQTK 121

QY 121 QCFERFYGGCLGNMNFETLECKNICEDGPNQFQVDNYGT 161

Db 122 QCFERFYGGCLGNMNFETLECKNICEDGPNQFQVDNYGT 162

RESULT 15

US-08-026-145-2

Sequence 2, Application US/08026145

Patent No. 5378614

GENERAL INFORMATION:

APPLICANT: Petersen, Jens G. Litske

APPLICANT: No. 5378614dfang, Ole Juul

TITLE OF INVENTION: Method for Making TFPI Analogues

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: No. 5378614o No. 5378614disk of No. 5378614th America, Inc.

STREET: 405 Lexington Avenue, Suite 6200

CITY: New York

STATE: N. Y.

COUNTRY: United States of America

ZIP: 10174-6201

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/026,145

FILING DATE: 19930302

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/828,920

FILING DATE: 27-JAN-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: WO PCT/DK/90/00212

FILING DATE: 17-AUG-1990

PRIOR APPLICATION DATA:

APPLICATION NUMBER: DK 4080/89

FILING DATE: 18-AUG-1989

ATTORNEY/AGENT INFORMATION:

NAME: Agriis, Cheryl H.

REGISTRATION NUMBER: 34086

REFERENCE/DOCKET NUMBER: 3321.214-US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212 867 0123

TELEFAX: 212 867 0298

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 304 amino acids

TYPE: AMINO ACID

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-026-145-2

Query Match 99.1%; Score 897; DB 1; Length 304;

Best Local Similarity 98.8%; Pred. No. 1.9e-91;

Matches 159; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 DSEDEHTIITDTELPPLKLMHSFCFAKADGPGCAIMKRFFNFTRQCEEFYGGCE 60

|||||

Db 29 DSEDEHTIITDELPLPLKLMHSFCAFKADDPCKAIMKRFFENIFRQCEEFIYGGCE 88
 Qy 61 GQNRFSLECKKMCRTDNANRIIKTTLOQEKPDFCFLEEDPGICRGYITRYFYNOOTK 120
 Db 89 GQNRFSLECKKMCRTDNANRIIKTTLOQEKPDFCFLEEDPGICRGYITRYFYNOOTK 148
 Qy 121 QCERFKYGGCLGNMNNFETLEECKNICEDGPNGFQVDNYGT 161
 Db 149 QCERFKYGGCLGNMNNFETLEECKNICEDGPNGFQVDNYGT 189

Search completed: December 2, 2002, 10:05:59
 Job time : 27.4328 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 2, 2002, 10:00:35 ; Search time 28.8358 Seconds
(without alignments)
536.751 Million cell updates/sec

Title: US-09-741-106-19

Perfect score: 905

Sequence: 1 DSEDEEHTIITDELPLK.....ECKNICEDGPGFQVDNYGT 161

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_73 :
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	897	99.1	304	1 TIHUGK	tissue factor path
2	836	92.4	304	1 JC2264	tissue factor path
3	653	72.2	300	2 S12443	lipoprotein-associ
4	645	71.3	299	2 I46937	tissue factor path
5	594	65.6	302	1 TIRTKG	tissue factor path
6	550	60.8	396	2 S53325	tissue factor path
7	274.5	30.3	235	2 A54951	tissue factor path
8	266.5	29.4	1558	2 C89114	protein C37C3.6a [
9	266.5	29.4	2167	2 T34395	hypothetical prote
10	257.5	28.5	2225	2 T26063	hypothetical prote
11	252	27.8	838	2 T20125	hypothetical prote
12	244	27.0	922	2 T23573	hypothetical prote
13	241.5	26.7	1416	2 E88550	protein zC84.6 [im
14	241.5	26.7	2844	2 S28291	hypothetical prote
15	240.5	26.6	1599	2 T16210	hypothetical prote
16	238	26.3	252	2 JG0185	hepatocyte growth
17	238	26.3	1965	2 T33216	hypothetical prote
18	235	26.0	1474	2 D88550	protein zC84.6 [im
19	233.5	25.8	352	1 TIB091	alpha-1-microglobu
20	233	25.7	1522	2 H88380	protein T22F7.3 [i
21	226.5	25.0	1203	2 T21275	hypothetical prote
22	221	24.4	1743	2 T26859	hypothetical prote
23	220.5	24.4	337	1 TIPGBI	alpha-1-microglobu
24	217	24.0	349	2 S35708	alpha-1-microglobu
25	216.5	23.9	123	2 A29652	inter-alpha-trypsi
26	211.5	23.4	125	1 TTH0B1	alpha-1-microglobu
27	210	23.2	349	2 S21089	alpha-1-microglobu
28	206.5	22.8	1391	2 T20406	hypothetical prote
29	203	22.4	1043	2 T19734	hypothetical prote

RESULT 1

TIHUGK

tissue factor pathway inhibitor precursor [validated] - human
N;Alternate names: extrinsic pathway inhibitor; lipoprotein-associated coagulation in
C;Species: Homo sapiens (man)
C;Date: 30-Jun-1992 #sequence.revision 30-Jun-1992 #text.change 08-Dec-2000
C;Accession: A23712; A39176; A28650; A60433; B60433; S13034; A34315; A38294; S03903
R;Girard, T.J.; Eddy, R.; Wesselschmidt, R.L.; MacPhail, L.A.; Likert, K.M.; Byers, M
J. Biol. Chem. 266, 5036-5041, 1991
A;Title: Structure of the human lipoprotein-associated coagulation inhibitor gene. In
A;Reference number: A23712; MUID:91161593; PMID:2002045
A;Accession: A23712
A;Molecule type: DNA
A;Residues: 1-304 <GIR>
A;Cross-references: GB:M59493; GB:M59499; NID:g187204; PIDN:AAA59526.1; PID:g187206
R;van der Logt, C.P.E.; Reitsma, P.H.; Bertina, R.M.
Biochemistry 30, 1571-1577, 1991
A;Title: Intron-exon organization of the human gene coding for the lipoprotein-associ
A;Reference number: A39176; MUID:91129227; PMID:1993173
A;Accession: A39176
A;Molecule type: DNA
A;Residues: 1-304 <VAN>
A;Cross-references: GB:M58650; GB:J05312; NID:g186827; PIDN:AAA59480.1; PID:g186829
R;Wun, T.C.; Kretzmer, K.K.; Girard, T.J.; Miletich, J.P.; Broze Jr., G.J.
J. Biol. Chem. 263, 6001-6004, 1988
A;Title: Cloning and characterization of a cDNA coding for the lipoprotein-associated
A;Reference number: A28650; MUID:88198127; PMID:2452157
A;Accession: A28650
A;Molecule type: mRNA
A;Residues: 1-304 <WUN>
A;Cross-references: GB:J03225; NID:g180545; PIDN:AAA52022.1; PID:g180546
A;Note: Part of this sequence, including the amino end of the mature protein, was con
R;Girard, T.J.; Warren, L.A.; Novotny, W.F.; Bimec, B.E.; Miletich, J.P.; Broze Jr.,
Thromb. Res. 55, 37-50, 1989
A;Title: Identification of the 1.4 KB and 4.0 KB messages for the lipoprotein associa
A;Reference number: A60433; MUID:89388722; PMID:2781520
A;Accession: A60433
A;Status: not compared with conceptual translation
A;Molecule type: protein
A;Residues: 'XX', 31-53, 'X', 55-56 <GI3>
A;Experimental source: recombinant material from mouse C137 cells
R;Girard, T.J.; McCourt, D.; Novotny, W.F.; MacPhail, L.A.; Likert, K.M.; Broze Jr.,
Biochem. J. 270, 621-625, 1990
A;Title: Endogenous phosphorylation of the lipoprotein-associated coagulation inhibit
A;Reference number: S13034; MUID:91054349; PMID:2122883
A;Accession: S13034
A;Molecule type: protein
A;Residues: 29-35 <GI4>

ALIGNMENTS

alpha-1-microglobu
gamma-1-microglobu
hypothetical prote
venom basic protei
hypothetical prote
venom basic protei
venom basic protei
taicotoxin serine
hypothetical prote
basic proteinase i
hypothetical prote
trypsin inhibitor,
venom basic protei
alpha-1-microglobu
venom basic protei
proteinase inhibit

R:Novotny, W.F.; Girard, T.J.; Milelich, J.P.; Broze Jr., G.J.
J. Biol. Chem. 264, 18832-18837, 1989
A:Title: Purification and characterization of the lipoprotein-associated coagulation inhibitor
A:Reference number: A34315; MUID:90036996; PMID:2553722
A:Accession: A34315
A:Molecule type: protein
A:Residues: 'XX', 31-33, 'L', 35-50 <NOV>
A:Experimental source: plasma
R:Pedersen, A.H.; Nordfang, O.; Norris, F.; Wiberg, F.C.; Christensen, P.M.; Moeller, K.
J. Biol. Chem. 265, 16786-16793, 1990
A:Title: Recombinant human extrinsic pathway inhibitor: Production, isolation, and characterization
A:Reference number: A38294; MUID:91009092; PMID:2211593
A:Accession: A38294
A:Molecule type: protein
A:Residues: 29-41 <PED>
R:Girard, T.J.; Warren, L.A.; Novotny, W.F.; Likert, K.M.; Brown, S.G.; Milelich, J.P.; Nature 338, 518-520, 1989
A:Title: Functional significance of the Kunitz-type inhibitory domains of lipoprotein-associated coagulation factor VIIa
A:Reference number: S03903; MUID:89181950; PMID:2927510
A:Contents: annotation: site-directed mutagenesis
C:Comment: The first Kunitz-type domain binds the factor VIIa/tissue factor complex; the C:Genetics:
A:Gene: GDB:TFPI
A:Cross-references: GDB:127364; OMIM:152310
A:Map position: 2932-2932
A:Introns: 41/1; 107/1; 120/1; 179/1; 210/1; 270/1
C:Function:
A:Description: regulates clotting by factor Xa-dependent inhibition of the coagulation factor Xa
A:Pathway: blood coagulation
C:Superfamily: tissue factor pathway inhibitor; animal Kunitz-type proteinase inhibitor
C:Keywords: anticoagulant; blood coagulation; duplication; glycoprotein; heparin binding
F:1-28/Domain: signal sequence #status predicted <SIG>
F:29-304/Product: tissue factor pathway inhibitor #status experimental <MAT>
F:125-175/Domain: animal Kunitz-type proteinase inhibitor homology <BP1>
F:217-267/Domain: animal Kunitz-type proteinase inhibitor homology <BP2>
F:284-289/Region: heparin binding #status predicted
F:30/Binding site: phosphate (Ser) (covalent) #status experimental
F:54-104,63-87,79-100,125-175,134-158,150-171,217-267,226-250,242-263/Disulfide bonds: #
F:64/Inhibitory site: Lys (coagulation factor VII/tissue factor complex) #status experimental
F:135/Inhibitory site: Arg (coagulation factor X) #status experimental
F:145,195,256/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:227/Inhibitory site: Arg (unidentified proteinase) #status predicted

Query Match 99.1%; Score 897; DB 1; Length 304;
Best Local Similarity 98.8%; Pred. No. 1.3e-76;
Matches 159; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 DSEDEEHTIITDTLPPLKLMHSCAFKADGPGPCRAIMKRFNFNFTQCEEFYGGCE 60
DB 29 DSEDEEHTIITDTLPPLKLMHSCAFKADGPGPCRAIMKRFNFNFTQCEEFYGGCE 88
QY 61 GNQRFESLECKKMTCDNANRIKTTLQEKPDPCFLEEDPGICRGYITRYFYNNQTK 120
DB 89 GNQRFESLECKKMTCDNANRIKTTLQEKPDPCFLEEDPGICRGYITRYFYNNQTK 148

RESULT 2
JC2264
tissue factor pathway inhibitor precursor - rhesus macaque
N:Alternate names: extrinsic pathway inhibitor; lipoprotein-associated coagulation inhibitor
C:Species: Macaca mulatta (rhesus macaque)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000
C:Accession: JC2264
R:Kamei, S.; Kamikubo, Y.; Hamuro, T.; Fujimoto, H.; Ishihara, M.; Yonemura, H.; Miyamoto, R.; Biochem. 115, 708-714, 1994
A:Title: Amino acid sequence and inhibitory activity of rhesus monkey tissue factor pathway inhibitor
A:Reference number: JC2264; MUID:94375417; PMID:8089087
A:Accession: JC2264

A:Molecule type: mRNA
A:Residues: 1-304 <KAM>
A:Cross-references: GB:S73337; NID:9685016; PIDN:AAB31955.1; PID:9685017
A:Experimental source: liver
C:Comment: This protein inhibits the activities of factor Xa and tissue factor-proteinase
C:Superfamily: tissue factor pathway inhibitor; animal Kunitz-type proteinase inhibitor
C:Keywords: anticoagulant; glycoprotein; serine proteinase inhibitor
F:1-28/Domain: signal sequence #status predicted <SIG>
F:29-304/Product: tissue factor pathway inhibitor #status predicted <MAT>
F:125-175/Domain: animal Kunitz-type proteinase inhibitor homology <BP1>
F:217-267/Domain: animal Kunitz-type proteinase inhibitor homology <BP2>
F:54-104,63-87,79-100,125-175,134-158,150-171,217-267,226-250,242-263/Disulfide bonds: #
F:64/Inhibitory site: Lys (coagulation factor VII/tissue factor complex) #status predicted
F:135/Inhibitory site: Arg (coagulation factor X) #status predicted
F:145,195,256/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:227/Inhibitory site: Arg (unidentified proteinase) #status predicted

Query Match 92.4%; Score 836; DB 1; Length 304;
Best Local Similarity 91.3%; Pred. No. 6.8e-71;
Matches 147; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

QY 1 DSEDEEHTIITDTLPPLKLMHSCAFKADGPGPCRAIMKRFNFNFTQCEEFYGGCE 60
DB 29 DSEDEEHTIITDTLPPLKLMHSCAFKADGPGPCRAIMKRFNFNFTQCEEFYGGCE 88
QY 61 GNQRFESLECKKMTCDNANRIKTTLQEKPDPCFLEEDPGICRGYITRYFYNNQTK 120
DB 89 GNQRFESLECKKMTCDNANRIKTTLQEKPDPCFLEEDPGICRGYITRYFYNNQTK 148

QY 121 QCERFYGGCLGNMNFTELECKNICEDGPGNFQVDNYGT 161
DB 149 QCERFYGGCLGNMNFTELECKNICEDGPGNFQVDNYGT 189

RESULT 3
S12143
lipoprotein-associated coagulation inhibitor precursor - rabbit
N:Alternate names: endothelial cell coagulation inhibitor; endothelial cell tissue factor
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 16-Jul-1999
C:Accession: S12143; A61373
R:Wesselschmidt, R.L.; Girard, T.J.; Broze Jr., G.J. Nucleic Acids Res. 18, 6440, 1990
A:Title: cDNA sequence of rabbit lipoprotein-associated coagulation inhibitor.
A:Reference number: S12143; MUID:91057146; PMID:2136251
A:Accession: S12143
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-300 <WES>
A:Cross-references: EMBL:X54708; NID:g1612; PIDN:CAA38515.1; PID:g1613
R:Colburn, P.; Crabb, J.W.; Buonassisi, V. J. Cell. Physiol. 148, 320-326, 1991
A:Title: Enhanced inhibition of tissue factor by the extended form of an endothelial
A:Reference number: A61373; MUID:91349227; PMID:1880157
A:Accession: A61373
A:Molecule type: protein
A:Residues: 25-33, 'X', 35-46 <COL>
C:Superfamily: tissue factor pathway inhibitor; animal Kunitz-type proteinase inhibitor
C:Keywords: anticoagulant; glycoprotein
F:50-100/Domain: animal Kunitz-type proteinase inhibitor homology <BP1>
F:121-171/Domain: animal Kunitz-type proteinase inhibitor homology <BP2>
F:213-263/Domain: animal Kunitz-type proteinase inhibitor homology <BP3>

Query Match 72.2%; Score 653; DB 2; Length 300;
Best Local Similarity 72.5%; Pred. No. 9.3e-54;
Matches 116; Conservative 13; Mismatches 31; Indels 0; Gaps 0;

QY 2 DSEDEEHTIITDTLPPLKLMHSCAFKADGPGPCRAIMKRFNFNFTQCEEFYGGCE 61
DB 26 AEEDEEFTNITDIKPLPQKPTSCAMKVDGPGPCRAYIKRFFNLAHQCEEFYGGCE 85
QY 62 GNQRFESLECKKMTCDNANRIKTTLQEKPDPCFLEEDPGICRGYITRYFYNNQTK 121

[illegible]

F:134/Inhibitory site: Arg (coagulation factor X) #status predicted
F:144,251,261/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:232/Inhibitory site: Lys (unidentified proteinase) #status predicted

Query Match 65.6%; Score 594; DB 1; Length 302;
Best Local Similarity 63.9%; Pred. No. 3.1e-48;
Matches 101; Conservative 26; Mismatches 31; Indels 0; Gaps 0;

QY 4 EDEBHTIITDTEPLPKLMHSFCAFKADGPCRAIMKRRFFNITRQCEBFIYGGCEGN 63
DB 31 EEDDDTINTDSELRPMKPLHTFCAMKAEDGPKAMIRSYFYFNWNSHQCEBFIYGGCEGN 90
QY 64 NRFSLECKKMCRTDRNANRIKTLTQEKPDPCFLEEDPGICRGYITRYFYNQTKQCE 123
DB 91 NRTDTLECRKTIPIGKYKTKITKTSAGKRPDPCFLEEDPGICRGFTRYFYNNQSKQCE 150
QY 124 RPKYGGCLGNMNPETLEECKNCEGDPNGFQVDNYGT 161
DB 151 QFKYGGCLGNMNPETLEECKNCEGDPVNEVQKGDYVT 188

RESULT 6
S53325
tissue factor pathway inhibitor - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 01-Aug-1995 #sequence_revision 01-Sep-1995 #text_change 16-Jul-1999
C:Accession: S53325
R:Girard, T. J.; Gailani, D.; Broze Jr., G. J.
Biochem. J. 303, 923-928, 1994
A:Title: Complementary DNA sequencing of canine tissue factor pathway inhibitor reveal
A:Reference number: S53325; MUID:95071310; PMID:7980463
A:Accession: S53325
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-396 <GIR>
C:Superfamily: animal Kunitz-type proteinase inhibitor homology
C:Keywords: serine proteinase inhibitor
F:53-103/Domain: animal Kunitz-type proteinase inhibitor homology <BP1>
F:125-175/Domain: animal Kunitz-type proteinase inhibitor homology <BP2>
F:309-359/Domain: animal Kunitz-type proteinase inhibitor homology <BP3>

Query Match 60.8%; Score 550; DB 2; Length 396;
Best Local Similarity 65.8%; Pred. No. 5.5e-44;
Matches 98; Conservative 24; Mismatches 21; Indels 6; Gaps 4;

QY 3 EDEBHTIITDTEPLPKLMHSFCAFKADGPCRAIMKRRFFNITRQCEBFIYGGCEGN 62
DB 31 DESEYPGITD-ELPPLRLHLSFCALKADGPCRAMIRNYFFNHTQOCEBFIYGGCEGN 89
QY 63 QNRFSLECKKMCRTDRNANRIKT-TLQO--EKPDFCFLEEDPGICRGYITRYFYNQOT 119
DB 90 QNRFSLECKEKCVRYPK--AKTLEKVLKRPDYCHMNEGSLGRGFTRYFYNNVS 147
QY 120 KQCRFYGGCLGNMNPETLEECKNICE 148
DB 148 SKCEGKYGGCLGNLNPETLEECKNTCE 176

RESULT 7
A54951
tissue factor pathway inhibitor-2 precursor - human
N:Alternate names: placental protein 5 (pp5)
C:Species: Homo sapiens (man)
C:Date: 31-May-1996 #sequence_revision 31-May-1996 #text_change 16-Jul-1999
C:Accession: A54951; I55185; A34029; C34029; B34029
R:Spreecher, C.A.; Kislai, W.; Mathewes, S.; Foster, D.C.
Proc. Natl. Acad. Sci. U.S.A. 91, 3353-3357, 1994
A:Title: Molecular cloning, expression, and partial characterization of a second human
A:Reference number: A54951; MUID:94211862; PMID:8159751
A:Accession: A54951
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-235 <RES>

A:Cross-references: GB:I27624; NID:g441149; PIDN:AAA20094.1; PID:g441150
A:Experimental source: placenta
R:Miyagi, Y.; Koshikawa, N.; Yasumitsu, H.; Miyagi, E.; Hirahara, F.; Aoki, I.; Misugi, J. Biochem. 116, 939-942, 1994
A:Title: cDNA cloning and mRNA expression of a serine proteinase inhibitor secreted by c
A:Reference number: I55185; MUID:95204397; PMID:7896752
A:Accession: I55185
A:Status: nucleic acid sequence not shown; translation not shown; translated from GB/EMBL
A:Molecule type: mRNA
A:Residues: 1-235 <RE2>
A:Cross-references: GB:D29992; NID:g484050; PIDN:BAA06272.1; PID:g484051
A:Note: parts of this sequence, including the amino end of the mature protein, were dete
R:Butzow, R.; Huhtala, M.L.; Bohn, H.; Virtanen, I.; Seppaelae, M.
Biochem. Biophys. Res. Commun. 150, 483-490, 1988
A:Title: Purification and characterization of placental protein 5.
A:Reference number: A34029; MUID:88106628; PMID:3276312
A:Accession: A34029
A:Molecule type: protein
A:Residues: 'A', 24-33, 'X', 35 <BU2>
A:Accession: C34029
A:Molecule type: protein
A:Residues: 47-50, 'X', 52-53 <BU2>
A:Accession: B34029
A:Molecule type: protein
A:Residues: 133, 'X', 135-137, 'X', 139-140, 'X', 142-144, 'X', 146 <BU3>
C:Genetics:
A:Gene: GDB:TFPI2
A:Cross-references: GDB:354485
C:Superfamily: tissue factor pathway inhibitor; animal Kunitz-type proteinase inhibitor
C:Keywords: anticoagulant; blood coagulation; duplication; glycoprotein; heparin binding
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-235/Product: tissue factor pathway inhibitor-2 #status predicted <MAT>
F:36-86/Domain: animal Kunitz-type proteinase inhibitor homology <BP1>
F:96-149/Domain: animal Kunitz-type proteinase inhibitor homology <BP2>
F:158-208/Domain: animal Kunitz-type proteinase inhibitor homology <BP3>
F:36-86, 45-69, 61-82, 96-149, 106-130, 122-145, 158-208, 167-191, 183-204/Disulfide bonds: #sta
Query Match 30.3%; Score 274.5; DB 2; Length 235;
Best Local Similarity 28.9%; Pred. No. 2e-18;
Matches 50; Conservative 29; Mismatches 43; Indels 51; Gaps 3;
Qy 26 CAFKADGPGCAIRKMKRFFNFTQCEFFYGGCEGNQNRFSLECKKMCNTR-DNANRI 84
Db 36 CLLPDVGPCALLRYYRYDQSCRFQYGGCEGNANFNFTWEACDDACWRIKVPKV 95
Qy 85 IK-----TTLOEK-----p 94
Db 96 CRLQVDDQCEGSTEKYFFNLSSMTCEKFFSGGCHNRNTRFPDEATCMGFCAPKKIP 155
Qy 95 DFCFLEEDPGICRITRYFYNOOTKOCERFKYGGCLGNMNFETLECKNIC 147
Db 156 SFQYSPKDEGLCSANVTYFNPRYTCDFTYTGCGGNDNNEVSRDCKRAC 208
RESULT 8
C89114
Protein C37C3.6a [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C:Accession: C89114
R:anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog
A:Reference number: A75000; MUID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C.ele
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A:Accession: C89114
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1558 <STO>
A:Cross-references: GB:chr_V; PIDN:AAC25867.1; PID:g3294501; GSPDB:GN00023; CESP:C37C3.6
C:Genetics:
A:Gene: C37C3.6a

A:Map position: 5

Query Match 29.4%; Score 266.5; DB 2; Length 1558;
Best Local Similarity 31.9%; Pred. No. 7.9e-17;
Matches 52; Conservative 23; Mismatches 55; Indels 33; Gaps 3;Qy 18 PLKLMHSFCAFKADGPGCAIRKMKRFFNFTQCEFFYGGCEGNQNRFSLECKKMC- 76
Db 1263 PQSMEDICRSQDAGPCETYSQDFYNFASQECETFTYGGGGLNFRSKDECEQRCF 1322

Qy 77 ----TRDNANR-----IKTTLOQKEP-----DFCFLEEDPG 104

Db 1323 FVHGAQPSAARQEQAPAAQAPQAPQASPVVPPQSSASPVVPSNRSKQRDACHLNVDQG 1382

Qy 105 ICRGYITRYFYNOOTKOCERFKYGGCLGNMNFETLECKNIC 147

Db 1383 RCKGAFDSWYEVATGSCVTFKTCGGGNANRFASKDQCESLC 1425

RESULT 9

T34395
hypothetical protein C37C3.6b - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 01-Dec-2000
C:Accession: T34395; T34394
R:Geisel, C.; Bradshaw, H.
submitted to the EMBL Data Library, July 1996
A:Description: The sequence of C. elegans cosmid C37C3.
A:Reference number: Z21518
A:Accession: T34395
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2167 <GEI>
A:Cross-references: EMBL:U64857; PIDN:AAC25867.1; GSPDB:GN00023; CESP:C37C3.6b
A:Experimental source: strain Bristol N2; clone C37C3
A:Accession: T34394
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1555, 'SKF' <GE2>
A:Cross-references: EMBL:U64857; PIDN:AAC25867.1; GSPDB:GN00023; CESP:C37C3.6a
C:Genetics:
A:Gene: CESP:C37C3.6b; CESP:C37C3.6a
A:Map position: 5
A:Introns: 32/3; 104/2; 156/2; 207/1; 459/2; 536/3; 577/2; 1105/3; 1367/1; 1438/1; 15Query Match 29.4%; Score 266.5; DB 2; Length 2167;
Best Local Similarity 31.9%; Pred. No. 1.1e-16;
Matches 52; Conservative 23; Mismatches 55; Indels 33; Gaps 3;Qy 18 PLKLMHSFCAFKADGPGCAIRKMKRFFNFTQCEFFYGGCEGNQNRFSLECKKMC- 76
Db 1263 PQSMEDICRSQDAGPCETYSQDFYNFASQECETFTYGGGGLNFRSKDECEQRCF 1322

Qy 77 ----TRDNANR-----IKTTLOQKEP-----DFCFLEEDPG 104

Db 1323 FVHGAQPSAARQEQAPAAQAPQAPQASPVVPPQSSASPVVPSNRSKQRDACHLNVDQG 1382

Qy 105 ICRGYITRYFYNOOTKOCERFKYGGCLGNMNFETLECKNIC 147

Db 1383 RCKGAFDSWYEVATGSCVTFKTCGGGNANRFASKDQCESLC 1425

RESULT 10

T26063
hypothetical protein W01F3.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T26063
R:Cummings, P.
submitted to the EMBL Data Library, March 1997
A:Reference number: Z20145
A:Accession: T26063

R:Thomas, K.
submitted to the EMBL Data Library, December 1992
A:Reference number: S28285
A:Accession: S28291
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-2844 <THO>
A:Cross-references: EMBL:Z19157

C:Genetics:
A:Introns: 14/1; 32/3; 57/1; 192/3; 277/1; 398/1; 439/1; 474/1; 497/1; 813/1; 1135/1; 12493/1; 2555/1; 2720/1; 2739/3; 2819/1
C:Superfamily: animal Kunitz-type proteinase inhibitor homology
F:220-274/Domain: animal Kunitz-type proteinase inhibitor homology <BP11>
F:343-395/Domain: animal Kunitz-type proteinase inhibitor homology <BP12>
F:442-492/Domain: animal Kunitz-type proteinase inhibitor homology <BP13>
F:546-598/Domain: animal Kunitz-type proteinase inhibitor homology <BP14>
F:654-706/Domain: animal Kunitz-type proteinase inhibitor homology <BP15>
F:1662-1716/Domain: animal Kunitz-type proteinase inhibitor homology <BP16>
F:1787-1839/Domain: animal Kunitz-type proteinase inhibitor homology <BP17>
F:1845-1895/Domain: animal Kunitz-type proteinase inhibitor homology <BP18>
F:1952-2004/Domain: animal Kunitz-type proteinase inhibitor homology <BP19>
F:2097-2152/Domain: animal Kunitz-type proteinase inhibitor homology <BP10>

Query Match 26.7%; Score 241.5; DB 2; Length 2844;
Best Local Similarity 33.1%; Pred. No. 3.2e-14;
Matches 56; Conservative 18; Mismatches 54; Indels 41; Gaps 5;
QY 24 SFCAFKADGPCRAIMKRRFFNIPTROCEFIYGGCEGNONRFESLECKKCMCTRDNRANR 83
Db 440 TICAQPLRIGDCTENVRKYWINARTROCFEYTGCGGNDNFDSDMCQNCFC--KNAIP 497
QY 84 IKTTLQOEKPDF-----CFLEEDPGI-C- 106
Db 498 EPKCIQGAQYKDMFGNEVTCNSMGCPANVBCYFDGSGWGCGCPKFTKFTCSLNTDSGIQCG 557
QY 107 RGYTRFYNOOTKOCERFKYGGCLGNMNNFETLECKNICEDG--PNG 153
Db 558 AGSTFKYYNPQTQNCESFOYNGCDGNSNNFANRDACESYCSVGGCPNG 606

RESULT 15
TI6210
hypothetical protein F30H5.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: TI6210
R:Pauley, A.; Stellyes, L.
submitted to the EMBL Data Library, June 1995
A:Description: The sequence of C. elegans cosmid F30H5.
A:Reference number: Z18478
A:Accession: TI6210
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1599 <PAU>
A:Cross-references: EMBL:U29096; NID:g861390; PID:g861393; PIDN:AAA68408.1; CESP:F30H5.3
A:Experimental source: strain Bristol N2
C:Genetics:
A:Gene: CESP:F30H5.3
A:Introns: 12/1; 59/2; 85/3; 124/3; 217/2; 534/3; 560/1; 1549/1

Query Match 26.6%; Score 240.5; DB 2; Length 1599;
Best Local Similarity 31.2%; Pred. No. 2.2e-14;
Matches 55; Conservative 21; Mismatches 49; Indels 51; Gaps 6;
QY 18 PLKLMHSFCAFKADGPCRAIMKRRFFNIPTROCEFIYGGCEGNONRFESLECKK--- 74
Db 566 PLRL-----GDCQSVRRYRYNAVTRACIFDYTGCGGNDNFDSDMCQNCFC 614
QY 75 -----MCTRDNRANRIKTTLQ-----QKPDPCFLEE 101
Db 615 NIIPQCQPGDAYKDYQGNYYVCSNSGAGNSCPVNYECYFDGYVWGCGCPKAYTCTLSP 674
QY 104-DPGI-C-RGYITRYFYNOOTKOCERFKYGGCLGNMNNFETLECKNICEDG--PNG 153

Db 675 HKGVTCGSGSSRYRYNQTQCEESYQYNGCDGNSNNFATREDCEGYCGVGGCPNG 730
Search completed: December 2, 2002, 10:04:56
Job time : 31.8358 secs

QY 121 QCFKFGGCLGNMNFETLECKNICDGPNGFQVDNYGT 161
|||||
Db 121 QCFKFGGCLGNMNFETLECKNICDGPNGFQVDNYGT 161
|||||
RESULT 2
US-09-766-778-1
; Sequence 1, Application US/09766778
; Patent No. US20010018204A1
; GENERAL INFORMATION:
; APPLICANT: Papathanassiou, Adonia E
; Green, Shawn J.
; TITLE OF INVENTION: Compositions and Methods for Inhibiting
; Cellular Proliferation
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.A.
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09766,778
; FILING DATE: 22-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/227,955
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Greene, Jamie L.
; REGISTRATION NUMBER: 32,467
; REFERENCE/DOCKET NUMBER: 05213-0290
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 818-3700
; TELEFAX: (404) 818-3799
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 276 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: Active-site
; LOCATION: 2..3
; OTHER INFORMATION: /note= "Site of partial
; phosphorylation"
; FEATURE:
; NAME/KEY: Active-site
; LOCATION: 117..118
; OTHER INFORMATION: /note= "Potential site for N-linked
; glycosylation"
; FEATURE:
; NAME/KEY: Active-site
; LOCATION: 167..168
; OTHER INFORMATION: /note= "Potential site for N-linked
; glycosylation"
; FEATURE:
; NAME/KEY: Active-site
; LOCATION: 228..229
; OTHER INFORMATION: /note= "Potential site for N-linked

; glycosylation"
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 26..76
; OTHER INFORMATION: /label= Kunitz-1
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 97..147
; OTHER INFORMATION: /label= Kunitz-2
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 189..239
; OTHER INFORMATION: /label= Kunitz-3
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-766-778-1
Query Match 99.1%; Score 897; DB 10; Length 276;
Best Local Similarity 98.8%; Pred. No. 1.4e-81;
Matches 159; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 DSEDEHTIITDTELPPLKLMHSFCFAKADGPGCAIMKRFFNFTQCCEFIYGCCE 60
|||||
Db 1 DSEDEHTIITDTELPPLKLMHSFCFAKADGPGCAIMKRFFNFTQCCEFIYGCCE 60
|||||
QY 61 GNORFESLECKKCMCTRDNANRIKTTLQOEKPDPCFLEEDPGICRGYITRYFYNNQTK 120
|||||
Db 61 GNORFESLECKKCMCTRDNANRIKTTLQOEKPDPCFLEEDPGICRGYITRYFYNNQTK 120
|||||
QY 121 QCFKFGGCLGNMNFETLECKNICDGPNGFQVDNYGT 161
|||||
Db 121 QCFKFGGCLGNMNFETLECKNICDGPNGFQVDNYGT 161
|||||
RESULT 3
US-09-957-607-47
; Sequence 47, Application US/09957607
; Patent No. US20020076728A1
; GENERAL INFORMATION:
; APPLICANT: MacLennan, John M.
; APPLICANT: Ladner, Robert C.
; TITLE OF INVENTION: Engineering Affinity Ligands for Macromolecules
; FILE REFERENCE: DIX-001.1 US-1
; CURRENT APPLICATION NUMBER: US/09/957,607
; CURRENT FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: 08/821,498
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: 08/619,885
; PRIOR FILING DATE: 1996-03-20
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 47
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-957-607-47
Query Match 36.1%; Score 327; DB 10; Length 58;
Best Local Similarity 98.3%; Pred. No. 4.8e-26;
Matches 57; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 22 MHSFCFAKADGPGCAIMKRFFNFTQCCEFIYGCCEGNORFESLECKKCMCTRD 79
|||||
Db 1 MHSFCFAKADGPGCAIMKRFFNFTQCCEFIYGCCEGNORFESLECKKCMCTRD 58
|||||
RESULT 4
US-09-957-607-48
; Sequence 48, Application US/09957607
; Patent No. US20020076728A1
; GENERAL INFORMATION:
; APPLICANT: MacLennan, John M.
; APPLICANT: Ladner, Robert C.
; TITLE OF INVENTION: Engineering Affinity Ligands for Macromolecules

FILE REFERENCE: DYX-001.1 US-1
CURRENT APPLICATION NUMBER: US/09/957,607
CURRENT FILING DATE: 2001-09-19
PRIOR APPLICATION NUMBER: 08/821,498
PRIOR FILING DATE: 1997-03-21
PRIOR APPLICATION NUMBER: 08/619,885
PRIOR FILING DATE: 1996-03-20
NUMBER OF SEQ ID NOS: 48
SOFTWARE: PatentIn version 3.1
SEQ ID NO 48
LENGTH: 58
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: amino acid sequence of LACI/F library
NAME/KEY: MISC.FEATURE
LOCATION: (13)-(13)
OTHER INFORMATION: X is P,H,L, or R
NAME/KEY: MISC.FEATURE
LOCATION: (16)-(16)
OTHER INFORMATION: X is A or G
NAME/KEY: MISC.FEATURE
LOCATION: (17)-(17)
OTHER INFORMATION: X is I,A,C,D,F,G,H,L,N,P,R,S,T,V or Y
NAME/KEY: MISC.FEATURE
LOCATION: (18)-(18)
OTHER INFORMATION: X is M,A,C,D,E,F,G,H,I,K,L,N,P,Q,R,S,T,V,W or Y
NAME/KEY: MISC.FEATURE
LOCATION: (19)-(19)
OTHER INFORMATION: X is K,A,E,G,L,M,P,Q,R,S,T,V or W
US-09-957-607-48

Query Match 32.8%; Score 297; DB 10; Length 58;
Best Local Similarity 89.7%; Pred. No. 4.4e-23;
Matches 52; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 22 MHSFCARADGPGCAIMKRRFFNFTQCCEEFYGGCGEGRNFRFESLECKMKCTRD 79
||||| 1: |||||
Db 1 MHSFCARADGPGCAIMKRRFFNFTQCCEEFYGGCGEGRNFRFESLECKMKCTRD 58

RESULT 5
US-09-827-948-5
Sequence 5, Application US/09827948
Patent No. US20010029034A1
GENERAL INFORMATION:
APPLICANT: Gentz, Reiner, L.
APPLICANT: Hsu, Tsu-An
APPLICANT: Ni, Jian
TITLE OF INVENTION: Tissue Factor Pathway Inhibitor-3
FILE REFERENCE: 1488.1290002
CURRENT APPLICATION NUMBER: US/09/827,948
CURRENT FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: US 09/013,896
PRIOR FILING DATE: 1998-01-27
NUMBER OF SEQ ID NOS: 31
SOFTWARE: PatentIn version 3.0
SEQ ID NO 5
LENGTH: 51
TYPE: PRT
ORGANISM: Human
US-09-827-948-5

Query Match 31.8%; Score 288; DB 10; Length 51;
Best Local Similarity 96.1%; Pred. No. 3e-22;
Matches 49; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 97 CFLEEDPGICGYYTRYFYNOTKOCERFKYGGCLGNMNNFETLECKNIC 147
||||| 1: |||||
Db 1 CFLEEDPGICGYYTRYFYNOTKOCERFKYGGCLGNMNNFETLECKNIC 51

RESULT 6
US-09-827-948-4
Sequence 4, Application US/09827948
Patent No. US20010029034A1
GENERAL INFORMATION:
APPLICANT: Gentz, Reiner, L.
APPLICANT: Hsu, Tsu-An
APPLICANT: Rosen, Craig A.
APPLICANT: Ni, Jian
TITLE OF INVENTION: Tissue Factor Pathway Inhibitor-3
FILE REFERENCE: 1488.1290002
CURRENT APPLICATION NUMBER: US/09/827,948
CURRENT FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: US 09/013,896
PRIOR FILING DATE: 1998-01-27
NUMBER OF SEQ ID NOS: 31
SOFTWARE: PatentIn version 3.0
SEQ ID NO 4
LENGTH: 51
TYPE: PRT
ORGANISM: Human
US-09-827-948-4

Query Match 31.3%; Score 283; DB 10; Length 51;
Best Local Similarity 96.1%; Pred. No. 9.3e-22;
Matches 49; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 26 CAFKADGPGCAIMKRRFFNFTQCCEEFYGGCGEGRNFRFESLECKKMC 76
||||| 1: |||||
Db 1 CAFKADGPGCAIMKRRFFNFTQCCEEFYGGCGEGRNFRFESLECKKMC 51

RESULT 7
US-10-086-176A-6
Sequence 6, Application US/10086176A
Patent No. US20020173465A1
GENERAL INFORMATION:
APPLICANT: Hembrough, Todd
APPLICANT: Papathanassiou, Adonia E.
APPLICANT: Green, Shawn J.
TITLE OF INVENTION: Compositions and Methods for Inhibiting Cellular Proliferation
FILE REFERENCE: 05213-0296 43170-266780
CURRENT APPLICATION NUMBER: US/10/086,176A
CURRENT FILING DATE: 2002-02-28
PRIOR APPLICATION NUMBER: US 09/766,778
PRIOR FILING DATE: 2001-01-22
PRIOR APPLICATION NUMBER: US 09/227,955
PRIOR FILING DATE: 1999-01-11
PRIOR APPLICATION NUMBER: US 08/796,850
PRIOR FILING DATE: 1997-02-06
PRIOR APPLICATION NUMBER: US 09/130,273
PRIOR FILING DATE: 1998-08-06
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn version 3.1
SEQ ID NO 6
LENGTH: 213
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic peptide
US-10-086-176A-6

Query Match 30.3%; Score 274.5; DB 9; Length 213;
Best Local Similarity 28.9%; Pred. No. 3e-20;
Matches 50; Conservative 29; Mismatches 43; Indels 51; Gaps 3;

QY 26 CAFKADGPGCAIMKRRFFNFTQCCEEFYGGCGEGRNFRFESLECKKMC 84
||||| 1: |||||
Db 14 CLPLDYGPCRALLLYYDRYQSCRFYGGCGEGRNANNFYWEACDCCWRIKPKV 73
QY 85 IK-----TTLOQEK-----P 94
::: ||

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 2, 2002, 09:59:55 : Search time 16.0199 Seconds
(without alignments)
416.837 Million cell updates/sec

Title: US-09-741-106-19

Perfect score: 905

Sequence: 1 DSEDEHTITDTPLPK.....ECKNICEDGNGFQVDNYGT 161

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	897	99.1	304	1	TFPI_HUMAN
2	836	92.4	304	1	TFPI_MACMU
3	658	72.7	300	1	TFPI_RABIT
4	594	65.6	302	1	TFPI_RAT
5	563	62.2	306	1	TFPI_MOUSE
6	274.5	30.3	235	1	TFP2_HUMAN
7	242	26.7	230	1	TFP2_MOUSE
8	241.5	26.7	1416	1	YN81_CABEL
9	233.5	25.8	352	1	AMBP_BOVIN
10	232	25.6	252	1	SPT2_HUMAN
11	231	25.5	252	1	SPT2_MOUSE
12	223	24.6	349	1	AMBP_MESAU
13	220.5	24.4	337	1	AMBP_PIG
14	217	24.0	349	1	AMBP_MOUSE
15	216.5	23.9	123	1	IATR_SHEEP
16	215.5	23.8	346	1	AMBP_MERUN
17	213.5	23.6	507	1	SPT1_MOUSE
18	212.5	23.5	513	1	SPT1_HUMAN
19	211.5	23.4	123	1	IATR_HORSE
20	210	23.2	349	1	AMBP_RAT
21	199.5	22.0	352	1	AMBP_HUMAN
22	192.5	21.3	355	1	AMBP_PLEPL
23	190	21.0	65	1	IVB3_VIPAA
24	188	20.8	69	1	CRPT_BOOMI
25	184	20.3	62	1	IVBT_ERIMA
26	182	20.1	61	1	IVB1_VIPAA
27	181	20.0	58	1	AXP2_ANTAF
28	173	19.1	134	1	EPPI_MOUSE
29	172	19.0	110	1	IBP_CARCR
30	171	18.9	67	1	IBPC_BOVIN
31	170	18.8	60	1	IVB2_DABRU
32	168.5	18.6	770	1	A4_RAT
33	167	18.5	57	1	IVB2_HEMHA

RESULT 1

TFPI_HUMAN ID TFPI_HUMAN STANDARD; PRT; 304 AA.
AC P10646; O95103; DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tissue factor pathway inhibitor precursor (TFPI) (Lipoprotein-
associated coagulation inhibitor) (LACI) (Extrinsic pathway inhibitor)
DE (BPI).
GN TFPI OR TFPI1 OR LACI.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID:9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM ALPHA).
RX MEDLINE=88198127; PubMed=2452157;
RA Wu T.-C., Kretzmer K.K., Girard T.J., Miletich J.P., Broze G.J. Jr.;
RT "Cloning and characterization of a cDNA coding for the lipoprotein-
associated coagulation inhibitor shows that it consists of three
tandem Kunitz-type inhibitory domains.";
RL J. Biol. Chem. 263:6001-6004(1988).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM ALPHA).
RX MEDLINE=91129227; PubMed=1933173;
RA van der Logt C.P.E., Reitsma P.H., Bertina R.M.;
RT "Intron-exon organization of the human gene coding for the
lipoprotein-associated coagulation inhibitor: the factor Xa dependent
inhibitor of the extrinsic pathway of coagulation.";
RL Biochemistry 30:1571-1577(1991).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM ALPHA).
RX MEDLINE=91161593; PubMed=2002045;
RA Girard T.J., Eddy R., Wesselschmidt R.L., Macphail L.A.,
RLikert K.M., Byers M.G., Shows T.B., Broze G.J. Jr.;
RT "Structure of the human lipoprotein-associated coagulation inhibitor
gene. Intro/exon gene organization and localization of the gene to
chromosome 2.";
RL J. Biol. Chem. 266:5036-5041(1991).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM ALPHA).
RX MEDLINE=89388722; PubMed=2781520;
RA Girard T.J., Warren L.A., Novotny W.F., Bejcek B.E., Miletich J.P.,
Broze G.J. Jr.;
RT "Identification of the 1.4 kb and 4.0 kb messages for the lipoprotein
associated coagulation inhibitor and expression of the encoded
protein.";
RL Thromb. Res. 55:37-50(1989).
RN [5]
RP SEQUENCE FROM N.A. (ISOFORM BETA).
RA Chang J.-Y., Monroe D.M., Roberts H.R.;
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A. (ISOFORM BETA).
RT TISSUE-Pancreas;

ALIGNMENTS

34 167 18.5 62 1 IP52_ANESU
35 167 18.5 770 1 A4_HUMAN
36 167 18.5 770 1 A4_MOUSE
37 166.5 18.4 265 1 TKD1_SHEEP
38 166 18.3 58 1 ISIK_HELPO
39 164 18.1 59 1 IVBI_DENAN
40 164 18.1 64 1 SPT3_HUMAN
41 161 17.8 57 1 SBPI_SARBU
42 160.5 17.7 164 1 TKD1_BOVIN
43 160 17.7 197 1 MCP1_MELCP
44 160 17.7 3137 1 CA36_CHICK
45 158 17.5 63 1 ICS3_BOMMO

P10280 anemonia su
P05067 homo sapien
P12023 mus musculu
Q29428 oviss aries
P00994 helix pomat
P00980 dendroaspis
P49223 homo sapien
P26228 sarcophaga
Q28201 bos taurus
P82968 melithaea c
P15989 gallus gall
P07481 bombyx mori

Strausberg R.;
Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
[7]
SEQUENCE OF 29-50.
MEDLINE=90036996; PubMed=2553722;
Novotny W.F., Girard T.J., Miletich J.P., Broze G.J. Jr.:
"Purification and characterization of the lipoprotein-associated
coagulation inhibitor from human plasma.";
J. Biol. Chem. 264:18832-18837(1989).
[8]
INHIBITORY SITES.
MEDLINE=89181950; PubMed=2927510;
Girard T.J., Warren L.A., Novotny W.F., Likert K.M., Brown S.G.,
Miletich J.P., Broze G.J. Jr.:
"Functional significance of the Kunitz-type inhibitory domains of
RT lipoprotein-associated coagulation inhibitor.";
Nature 338:518-520(1989).
[9]
CARBOHYDRATE-LINKAGE SITES.
MEDLINE=96224851; PubMed=8639592;
Nakanahara Y., Miyata T., Hamuro T., Funatsu A., Miyagi M.,
Tsunashima S., Kato H.:
"Amino acid sequence and carbohydrate structure of a recombinant
human tissue factor pathway inhibitor expressed in Chinese hamster
ovary cells: one N-and two O-linked carbohydrate chains are located
between Kunitz domains 2 and 3 and one N-linked carbohydrate chain is
in Kunitz domain 2";
Biochemistry 35:6450-6459(1996).
[10]
REVIEW.
MEDLINE=91104709; PubMed=2271516;
Broze G.J. Jr., Girard T.J., Novotny W.F.:
"Regulation of coagulation by a multivalent Kunitz-type inhibitor.";
Biochemistry 29:7539-7546(1990).
[11]
STRUCTURE BY NMR OF 121-182.
MEDLINE=97342711; PubMed=9199408;
Burgering M.J., Orbons L.P., van der Doelen A., Mulders J.,
Theunissen H.J., Grootenhuys P.D., Bode W., Huber R., Stubbs M.T.:
"The second Kunitz domain of human tissue factor pathway inhibitor:
cloning, structure determination and interaction with factor Xa";
J. Mol. Biol. 269:395-407(1997).
CC -1- FUNCTION: INHIBITS FACTOR X (X(A)) DIRECTLY AND, IN A XA-DEPENDENT
WAY, INHIBITS VII(A)/TISSUE FACTOR ACTIVITY, PRESUMABLY BY FORMING
A QUATERNARY X(A)/LAGI/VII(A)/TF COMPLEX. IT POSSESSES AN
ANTI THROMBOTIC ACTION AND ALSO THE ABILITY TO ASSOCIATE WITH
LIPOPROTEINS IN PLASMA.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; alpha/'FP'alpha (shown here) and
beta/'FP'beta; are produced by alternative splicing.
CC -1- TISSUE SPECIFICITY: MOSTLY IN ENDOTHELIAL CELLS.
CC -1- DOMAIN: THIS INHIBITOR CONTAINS THREE INHIBITORY DOMAINS.
CC -1- PTM: O-GLYCOSYLATED.
CC -1- SIMILARITY: CONTAINS 3 BPTI/KUNITZ INHIBITOR DOMAINS.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

EMBL; J03225; AAA52022.1; -.
DR EMBL; M58650; AAA59480.1; -.
DR EMBL; M58644; AAA59480.1; JOINED.
DR EMBL; M58645; AAA59480.1; JOINED.
DR EMBL; M58646; AAA59480.1; JOINED.
DR EMBL; M58647; AAA59480.1; JOINED.
DR EMBL; M58648; AAA59480.1; JOINED.
DR EMBL; M58649; AAA59480.1; JOINED.
DR EMBL; M59499; AAA59526.1; -.
DR EMBL; J559493; AAA59526.1; JOINED.

RL J. Biochem. 116:939-942(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=94211862; PubMed=8159751;
RA Sprecher C.A., Kiesel W., Mathewes S., Foster D.C.;
RT "Molecular cloning, expression, and partial characterization of a
RT second human tissue-factor-pathway inhibitor.";
RL Biochim. Biophys. Res. Commun. 151:3353-3357(1994).
RN [3]
RP SEQUENCE FROM N.A.
RA Maggii L.;
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RX PubMed=11342222;
RA Kamei S., Kazama Y., Kuiper J.L., Foster D.C., Kiesel W.;
RT "Genomic structure and promoter activity of the human tissue factor
RT pathway inhibitor-2 gene.";
RL Biochim. Biophys. Res. Commun. 151:430-435(2001).
RN [5]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
RN [6]
RP PARTIAL SEQUENCE OF 23-35; 47-53 AND 133-146.
RC TISSUE=Placenta;
RX MEDLINE=88106628; PubMed=3276312;
RA Buetzow R., Huhtala M.-L., Bohn H., Virtanen I., Seppaelae M.;
RT "Purification and characterization of placental protein 5.";
RL Biochem. Biophys. Res. Commun. 150:483-490(1988).
RN [7]
RP ERATUM.
RA Buetzow R., Huhtala M.-L., Bohn H., Virtanen I., Seppaelae M.;
RL Biochem. Biophys. Res. Commun. 151:630-631(1988).
CC -!- FUNCTION: SEEMS TO INHIBIT TRYPSIN, FACTOR VII(A)/TISSUE FACTOR,
CC WEAKLY FACTOR XA. HAS NO EFFECT ON THROMBIN.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: UMBILICAL VEIN ENDOTHELIAL CELLS, LIVER,
CC PLACENTA, HEART, PANCREAS, AND MATERNAL SERUM AT ADVANCED
CC PREGNANCY.
CC -!- DOMAIN: THIS INHIBITOR CONTAINS THREE INHIBITORY DOMAINS.
CC -!- SIMILARITY: CONTAINS 3 BPTI/KUNITZ INHIBITOR DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D29992; BAA06272.1; -
DR EMBL; L27624; ARA20094.1; -
DR EMBL; AC002076; AAB54049.1; -
DR EMBL; AF217542; AAK13254.1; -
DR EMBL; BC005330; AAB05330.1; -
DR PIR; A34029; A34029.
DR PIR; B34029; B34029.
DR PIR; C34029; C34029.
DR HSP; F12111; 1KNT.
DR Genew; HGNC:11761; TFP2.
DR MIM; 600033; -
DR InterPro; IPR002223; Kunitz_BPTI.
DR Pfam; PF00014; Kunitz_BPTI; 3.
DR PRINTS; PR00759; BASICPTASE.
DR ProDom; PD000222; Kunitz_BPTI; 3.
DR SMART; SM00131; KU; 2.
DR PROSITE; PS00280; BPTI_KUNITZ_1; 2.
DR PROSITE; PS02079; BPTI_KUNITZ_2; 3.
KW Serine protease inhibitor; Glycoprotein; Repeat; Signal;
KW Blood coagulation; Polymorphism.

FT SIGNAL 1 22
FT CHAIN 23 235
FT DOMAIN 36 86
FT DOMAIN 96 149
FT DOMAIN 158 208
FT DOMAIN 213 217
FT ACT_SITE 46 47
FT ACT_SITE 107 108
FT ACT_SITE 168 169
FT DISULFID 36 86
FT DISULFID 45 69
FT DISULFID 61 82
FT DISULFID 96 149
FT DISULFID 106 130
FT DISULFID 122 145
FT DISULFID 158 208
FT DISULFID 167 191
FT DISULFID 183 204
FT CARBOHYD 116 116
FT CARBOHYD 170 170
FT VARIANT 102 102
FT CONFLICT 23 23
FT SEQUENCE 235 AA; 26934 MW; 975ABA5C53F7C65F CRC64;
SQ
Query Match 30.3%; Score 274.5; DB 1; Length 235;
Best Local Similarity 28.9%; Pred. No. 2.4e-19;
Matches 50; Conservative 29; Mismatches 43; Indels 51; Gaps 3;
QY 26 CAFKADGFCRAIMKRFNFNIFTRQCEEFYGGCEGNQNFESLEBCKMKCTR-DNANEI 84
DB 36 CLLPDYGPCRALLLRYDYDRYQSCRFYGGCEGNANFNFTWEACDDACWRKVPKV 95
QY 85 IK-----TTLOQEK-----P 94
DB 96 CRLOVSDDCCESTEKYEFLNLSMTCEKFFSGGCHRNLENFPDEATCMGCPKPIP 155
QY 95 DRFLLEDPGICRGYTRYFYNOOTCRKFRKYGGCLGNMNNFETLECKNIC 147
DB 156 SFCYSPKDEGLCSANVTRYFNRYTCTDAFTYTGCGGNDNNFVSREDCKRAC 208
RESULT 7
TFP2_MOUSE
ID TFP2_MOUSE STANDARD; PRT; 230 AA.
AC Q35536;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Tissue factor pathway inhibitor 2 precursor (TFPI-2).
GN TFP12.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ICR; TISSUE=Placenta;
RX MEDLINE=97101108; PubMed=8945635;
RA Miyagi Y., Yasumitsu H., Mizushima H., Koshikawa N., Matsuda Y.,
RA Itoh H., Hori T., Aoki I., Misugi K., Miyazaki K.;
RT "Cloning of the cDNA encoding mouse PF5/TFPI-2 and mapping of the gene
RT to chromosome 6.";
RL DNA Cell Biol. 15:947-954(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SvJ;
RA Kazama Y., Kamei S., Kuiper J., Foster D.C., Kiesel W.;
RT "Nucleotide sequence of the gene encoding murine tissue factor pathway
RL inhibitor-2.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: SEEMS TO INHIBIT TRYPSIN, FACTOR VII(A)/TISSUE FACTOR,
CC WEAKLY FACTOR XA. HAS NO EFFECT ON THROMBIN.

CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN PLACENTA. ALSO EXPRESSED
CC IN LIVER AND KIDNEY.
CC -!- DOMAIN: THIS INHIBITOR CONTAINS THREE INHIBITORY DOMAINS.
CC -!- SIMILARITY: CONTAINS 3 BPTI/KUNITZ INHIBITOR DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: D50586; BAA22585.1; .
CC EMBL: AF180353; AAF40412.1; .
CC HSSP: P12111; 1KNT.
CC MGD: MGI:108543; Tfp12.
CC InterPro: IPR002223; Kunitz_BPTI.
CC Pfam: PF00014; Kunitz_BPTI; 3.
CC PRINTS: PR00759; BASICPTASE.
CC ProDom: PD000222; Kunitz_BPTI; 2.
CC SMART: SM00131; KU; 2.
CC PROSITE: PS00280; BPTI_KUNITZ_1; 2.
CC PROSITE: PS0279; BPTI_KUNITZ_2; 3.
CC Serine protease inhibitor; Glycoprotein; Repeat; Signal;
KW Blood coagulation.
FT SIGNAL 1 22 BY SIMILARITY.
FT CHAIN 23 230 TISSUE FACTOR PATHWAY INHIBITOR 2.
FT DOMAIN 36 86 BPTI/KUNITZ INHIBITOR 1.
FT DOMAIN 96 146 BPTI/KUNITZ INHIBITOR 2.
FT DOMAIN 156 206 BPTI/KUNITZ INHIBITOR 3.
FT ACT_SITE 46 47 REACTIVE BOND (BY SIMILARITY).
FT ACT_SITE 106 107 REACTIVE BOND (BY SIMILARITY).
FT ACT_SITE 166 167 REACTIVE BOND (BY SIMILARITY).
FT DISULFID 36 86 BY SIMILARITY.
FT DISULFID 45 69 BY SIMILARITY.
FT DISULFID 61 82 BY SIMILARITY.
FT DISULFID 96 146 BY SIMILARITY.
FT DISULFID 105 129 BY SIMILARITY.
FT DISULFID 121 142 BY SIMILARITY.
FT DISULFID 136 206 BY SIMILARITY.
FT DISULFID 165 189 BY SIMILARITY.
FT DISULFID 181 202 BY SIMILARITY.
FT CARBOHYD 168 168 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 178 178 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT SEQUENCE 230 AA; 26137 MW; 57EADB2E36521C7B CRC64;

Query Match 26.7%; Score 242; DB 1; Length 230;
Best Local Similarity 26.9%; Pred. No. 3e-16;
Matches 50; Conservative 27; Mismatches 57; Indels 52; Gaps 4;
QY 26 CAFKADGDCRAIMKRFNFNFTROCEFIYGCCEGNQNFESLEECKMCTR----- 78
Db 36 CLLPLDAGPCQALIPKFIYDRQCKRFRNYGCGLGNNFNHSDLCQCTGCSIEKVPV 95
QY 79 -----DNAN-----RIAKTTLQKQK-----PDF 96
Db 96 CRSELKTYCDKRNFRFFNLTMTCEPLRPLGLCSRTNVFSEATCKGLCEPKHIFSF 155
QY 97 CFLLEDPGICRGYITRYFYNQTKQCERFYGGCLGNMNNFTLECKNICDGG---PNG 153
Db 156 CSSPKDEGLCSANVTRFYFNSRNKTCETFTYTCGCGNENNFYLDACHRACVKGWKKPKR 215
QY 154 FQVDNY 159
Db 216 WKIGDF 221

RESULT 8
YNS1_CAEEL STANDARD; PRT; 1416 AA.
ID YNS1_CAEEL
AC Q03610;

DT 01-FEB-1994 (Rel. 28, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein ZC84.1 in chromosome III.
GN ZC84.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,
RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,
RA Sims M., Smaldon N., Smith A., Smith M., Sonhammer E., Staden R.,
RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,
RA Wohlman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL Nature 368:32-38(1994).
RN [2]
RP REVISIONS.
RA Jones S.J.M.;
RA Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: STRONG, TO D104.3.
CC -!- SIMILARITY: CONTAINS 5 BPTI/KUNITZ INHIBITOR DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: Z19157; CAA79569.1; .
CC PIR: S28291; S28291.
CC HSSP: P00974; 1PAK.
CC WormPep: ZC84.1; CE15020.
CC InterPro: IPR000561; EGF-like.
CC InterPro: IPR002223; Kunitz_BPTI.
CC InterPro: IPR002899; WRI/EB.
CC Pfam: PF00014; Kunitz_BPTI; 5.
CC Pfam: PF01683; EB; 3.
CC PRINTS: PR00759; BASICPTASE.
CC ProDom: PD000222; Kunitz_BPTI; 5.
CC SMART: SM00131; KU; 5.
CC SMART: SM00289; WRI; 13.
CC PROSITE: PS00280; BPTI_KUNITZ_1; 3.
CC PROSITE: PS0279; BPTI_KUNITZ_2; 5.
CC PROSITE: PS01186; EGF_2; UNKNOWN 1.
KW Hypothetical protein; Serine protease inhibitor; Repeat.
FT DOMAIN 212 266 BPTI/KUNITZ INHIBITOR 1.
FT DOMAIN 337 387 BPTI/KUNITZ INHIBITOR 2.
FT DOMAIN 434 484 BPTI/KUNITZ INHIBITOR 3.
FT DOMAIN 538 590 BPTI/KUNITZ INHIBITOR 4.
FT DOMAIN 646 698 BPTI/KUNITZ INHIBITOR 5.
FT SEQUENCE 1416 AA; 152986 MW; 531CACE1CB2F70D CRC64;
Query Match 26.7%; Score 241.5; DB 1; Length 1416;
Best Local Similarity 33.1%; Pred. No. 2.2e-15;
Matches 56; Conservative 18; Mismatches 54; Indels 41; Gaps 5;

QY 24 SFCAFKADGDCRAIMKRFNFNFTROCEFIYGCCEGNQNFESLEECKMCTRDNANR 83
Db 432 TICAQLRIGDCTENVKRYWYNARTQCMFEYTCGCGNNDNFSDIMDCNFC--KNAIP 489

QY 84 IKTTLQKPKDF-----CFLEEDPGI-C- 106
Db 490 EPKCIQOAYKDFGNFVTCNSMGCPANYECYFDGSQWGCCTPKAFTCSLNTDSGIQCG 549
QY 107 RGYITRYFYNQTKQGERPKYGCGLGNMNNFTLEECNIBCDG--PNG 153
Db 550 AGSTERYYYNPQTQNCSPQYNGCDGNSNFRNADACESYCVSGVCGPNG 598

RESULT 9
AMBP_BOVIN STANDARD; PRT: 352 AA.
ID AMBP_BOVIN P35420; Q28020;
AC P00378; P35420; Q28020;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE AMBP protein precursor (Contains: Alpha-1-microglobulin; Inter-alpha-
trypsin inhibitor light chain (ITI-IC) (Bikunin) (HI-30) (BI-14)
DE (Cumulus extracellular matrix stabilizing factor) (ESF)).
GN AMBP OR ITIL.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OC NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=96201710; PubMed=8611630;
RA Lindqvist A., Akerstrom B.;
RA "Bovine alpha 1-microglobulin/bikunin. Isolation and characterization
of liver cDNA and urinary alpha 1-microglobulin.";
RL Biochim. Biophys. Acta 1306:98-106(1996).
RN [2]
RP SEQUENCE OF 227-349.
RX MEDLINE=85225967; PubMed=2408637;
RA Hochstrasser K., Wächter E., Albrecht G.J., Reisinger P.;
RT "Kunitz-type proteinase inhibitors derived by limited proteolysis of
the inter-alpha-trypsin inhibitor, X. The amino-acid sequences of the
trypsin-released inhibitors from horse and pig inter-alpha-trypsin
inhibitors.";
RT Biol. Chem. Hoppe-Seyler 366:473-478(1985).
RN [3]
RP SEQUENCE OF 227-348.
RX MEDLINE=84133807; PubMed=6199275;
RA Hochstrasser K., Wächter E.;
RT "Kunitz-type proteinase inhibitors derived by limited proteolysis of
the inter-alpha-trypsin inhibitor, VII. Determination of the
amino-acid sequence of the trypsin-released inhibitor from bovine
inter-alpha-trypsin inhibitor.";
RL Hoppe-Seyler's Z. Physiol. Chem. 364:1679-1687(1983).
RN [4]
RP SEQUENCE OF 206-219.
RC TISSUE=Petal serum;
RX MEDLINE=92291130; PubMed=1376324;
RA Chen L., Mao S.J., Larsen W.J.;
RT "Identification of a factor in fetal bovine serum that stabilizes the
cumulus extracellular matrix. A role for a member of the inter-alpha-
trypsin inhibitor family.";
RL J. Biol. Chem. 267:12380-12386(1992).
RN [5]
RP REACTIVE SITES.
RX MEDLINE=84133808; PubMed=6199276;
RA Hochstrasser K., Albrecht G.J., Schoenberger O.L., Wächter E.;
RT "Kunitz-type proteinase inhibitors derived by limited proteolysis of
the inter-alpha-trypsin inhibitor, VII. Characterization of the
bovine inhibitor as double-headed trypsin-elastase inhibitor.";
RL Hoppe-Seyler's Z. Physiol. Chem. 364:1689-1696(1983).
CC -!- FUNCTION: ALPHA-1-MICROGLOBULIN OCCURS IN MANY PHYSIOLOGICAL
CC FLUIDS INCLUDING PLASMA, URINE, AND CEREBROSPINAL FLUID. IT
CC APPEARS NOT ONLY AS A FREE MONOMER BUT ALSO IN COMPLEXES WITH IGA
CC AND ALBUMIN.

CC CC -!- FUNCTION: INTER-ALPHA-TRYPSIN INHIBITOR, PRESENT IN PLASMA AND
CC URINE, INHIBITS TRYPSIN, PLASMIN, AND LYSOSOMAL GRANULOCYTIC
CC ELASTASE.
CC CC -!- FUNCTION: MAY DIFFUSE INTO FOLLICULAR FLUID AFTER AN OVULATORY
CC STIMULUS TO ACT AS STRUCTURAL LINKER THAT ENSURE NORMAL CUMULUS
CC EXPANSION, THROUGH STABILIZATION OF THE CUMULUS EXTRACELLULAR
CC MATRIX THUS SUPPORTING THE PROCESS OF OVULATION.
CC CC -!- SUBUNIT: I-ALPHA-I plasma protease inhibitors are assembled from
CC one or two heavy chains (H1, H2 or H3) and one light chain,
CC bikunin. Inter-alpha-inhibitor (I-ALPHA-I) is composed of H1, H2
CC and bikunin, inter-alpha-like inhibitor (I-ALPHA-LI) of H2 and
CC bikunin, and pre-alpha-inhibitor (P-ALPHA-I) of H3 and bikunin (By
CC similarity).
CC CC -!- SUBCELLULAR LOCATION: Secreted.
CC CC -!- TISSUE SPECIFICITY: Synthesized in the liver and secreted in
CC plasma.
CC CC -!- PTM: The precursor is proteolytically processed into two
CC separately functioning proteins.
CC CC -!- PTM: Alpha-1-microglobulin contains covalently linked brown-
CC yellow chromophores (By similarity).
CC CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE LIPOCALIN
CC FAMILY.
CC CC -!- SIMILARITY: CONTAINS 2 BPTI/KUNITZ INHIBITOR DOMAINS.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U35642; AAB07599.1; -.
CC PIR; A01209; TIBOBI.
CC HSP; P02760; IBIK.
CC InterPro; IPR002223; Kunitz_BPTI.
CC InterPro; IPR002345; Lipocalin.
CC Pfam; PF00014; Kunitz_BPTI; 2.
CC Pfam; PF00061; Lipocalin; 1.
CC PRINTS; PR00759; BASICPTASE.
CC PRINTS; PR00179; LIPOCALIN.
CC ProDom; PD000222; Kunitz_BPTI; 2.
CC SMART; SM00131; KU; 2.
CC PROSITE; PS00280; BPTI_KUNITZ_1; 2.
CC PROSITE; PS0279; BPTI_KUNITZ_2; 2.
CC PROSITE; PS00213; LIPOCALIN; 1.
CC Glycoprotein; Plasma; Signal; Serine protease inhibitor; Repeat;
CC Lipocalin.
FT SIGNAL 1 19 BY SIMILARITY.
FT CHAIN 20 203 ALPHA-1 MICROGLOBULIN.
FT CHAIN 206 352 INTER-ALPHA-TRYPSIN INHIBITOR LIGHT
FT CHAIN CHAIN.
FT DOMAIN 231 281 BPTI/KUNITZ INHIBITOR 1.
FT DOMAIN 287 337 BPTI/KUNITZ INHIBITOR 2.
FT BINDING 53 53 CHROMOPHORE (BY SIMILARITY).
FT BINDING 111 111 CHROMOPHORE (BY SIMILARITY).
FT BINDING 137 137 CHROMOPHORE (BY SIMILARITY).
FT BINDING 149 149 CHROMOPHORE (BY SIMILARITY).
FT DISULFID 91 188 BY SIMILARITY.
FT DISULFID 231 281 BY SIMILARITY.
FT DISULFID 240 264 BY SIMILARITY.
FT DISULFID 256 277 BY SIMILARITY.
FT DISULFID 287 337 BY SIMILARITY.
FT DISULFID 296 320 BY SIMILARITY.
FT DISULFID 312 333 BY SIMILARITY.
FT ACT_SITE 241 242 INHIBITORY SITE (P1) (CHYMOTRYPSIN,
FT ELASTASE).
FT ACT_SITE 297 298 INHIBITORY SITE (P1) (TRYPSIN).
FT CARBOHYD 115 115 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 223 223 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 250 250 N-LINKED (GLCNAC. . .).
FT CONFLICT 209 209 T -> G (IN REF. 4).

```
FT CONFLICT 217 217 A -> D (IN REF. 4).
FT CONFLICT 268 268 G -> L (IN REF. 2 AND 3).
FT CONFLICT 274 274 E -> O (IN REF. 2 AND 3).
FT CONFLICT 298 298 SY -> AF (IN REF. 2 AND 3).
FT CONFLICT 330 330 E -> Q (IN REF. 2 AND 3).
FT CONFLICT 346 346 E -> R (IN REF. 2 AND 3).
SQ SEQUENCE 352 AA; 39235 MW; ED31C5CA02E70B19 CRC64;

Query Match 25.8%; Score 233.5; DB 1; Length 352;
Best Local Similarity 37.7%; Pred. No. 3.1e-15;
Matches 46; Conservative 13; Mismatches 48; Indels 15; Gaps 1;

Qy 26 CAFKADGDCRAIMKREFFNIIFTRQCEFIYGGCEGNQNRFSLECKMKCTRDNANRII 85
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 231 COLDYSQPCGLGKRYFYNGTSACETFLYGGCMGNFSLSEKCLQTCRTVEA---- 286

Qy 86 KTTLOQKPDCEFLPEDPGICRGYITRYFYNQTKQCEKRYGCGCLGNMNFETLECKN 145
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 287 -----CNLPVGGPCRSYQLWAFDAVKGKCVRFYGGCKGNKGFYSEKCKE 335

Qy 146 IC 147
|
Db 336 YC 337

RESULT 10
SPT2_HUMAN STANDARD; PRT; 252 AA.
AC O43291; O00271; O14895; Q969E0;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Kunitz-type protease inhibitor 2 precursor (Hepatocyte growth factor
DE activator inhibitor type 2) (HAI-2) (Placental bikunin).
GN SPINT2 OR HAI2 OR KOP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID:9606;
[1]
RN Kunitz-type protease inhibitor 2.
RP SEQUENCE FROM N.A.
RX MEDLINE=98010584; PubMed=9346890;
RA Kawaguchi T., Qin L., Shimomura T., Kondo J., Matsumoto K., Denda K.,
RA Kitamura N.;
RT "Purification and cloning of hepatocyte growth factor activator
RT inhibitor type 2, a Kunitz-type serine protease inhibitor.";
RL J. Biol. Chem. 272:27558-27564(1997).
[2]
RN Kunitz-type protease inhibitor 2.
RP SEQUENCE FROM N.A., AND SEQUENCE OF 28-74.
RC TISSUE=Placenta;
RX MEDLINE=97277372; PubMed=9115294;
RA Marlor C.W., Delaria K.A., Davis G., Muller D.K., Greve J.M.,
RA Tamburini P.;
RT "Identification and cloning of human placental bikunin, a novel serine
RT protease inhibitor containing two Kunitz domains.";
RL J. Biol. Chem. 272:12202-12208(1997).
[3]
RN Kunitz-type protease inhibitor 2.
RP SEQUENCE FROM N.A.
RC TISSUE=Pancratic cancer;
RX MEDLINE=98094245; PubMed=9434156;
RA Mueller-Pillasch F., Wallrapp C., Bartels K., Varga G., Friess H.,
RA Buechler M., Adler G., Gress T.M.;
RT "Cloning of a new Kunitz-type protease inhibitor with a putative
RT transmembrane domain overexpressed in pancreatic cancer.";
RL Biochim. Biophys. Acta 1395:88-95(1998).
[4]
RN Kunitz-type protease inhibitor 2.
RP SEQUENCE FROM N.A., AND VARIANT LEU-200.
RC TISSUE=Colon, and Ovary;
RA Strausberg R.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: INHIBITOR OF HGF ACTIVATOR. ALSO INHIBITS PLASMIN,
CC PLASMA AND TISSUE KALLIKREIN, AND FACTOR XIA.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
```

```
CC -!- TISSUE SPECIFICITY: EXPRESSED IN PLACENTA, KIDNEY, PANCREAS,
CC PROSTATE, TESTIS, THYMUS, AND TRACHEA.
CC -!- DOMAIN: THIS INHIBITOR CONTAINS TWO INHIBITORY DOMAINS.
CC -!- SIMILARITY: CONTAINS 2 BPTI/KUNITZ INHIBITOR DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB006534; BAA25024.1; -
CC EMBL; U78095; AAC02781.1; -
CC EMBL; AF027205; AAB84031.1; -
CC EMBL; BC001668; AAH01668.1; -
CC EMBL; BC007705; AAH07705.1; -
CC EMBL; BC011951; AAH11951.1; -
CC EMBL; BC011955; AAH11955.1; -
CC EMBL; BC012868; AAH12868.1; -
CC HSP; P05067; IAAIP.
CC Genew; HGNC:11247; SPINT2.
CC MIM; 605124; -
CC InterPro; IPR002223; Kunitz_BPTI.
CC Pfam; PF00014; Kunitz_BPTI; 2.
CC PRINTS; PR00759; BASICPTASE.
CC ProDom; PD000222; Kunitz_BPTI; 2.
CC SMART; SM00131; KU; 2.
CC PROSITE; PS00280; BPTI_KUNITZ_1; 2.
CC PROSITE; PS0279; BPTI_KUNITZ_2; 2.
KW Serine protease inhibitor; Repeat; Glycoprotein; Transmembrane;
KW Signal; Polymorphism.
FT SIGNAL 1 27
FT CHAIN 28 252 KUNITZ-TYPE PROTEASE INHIBITOR 2.
FT DOMAIN 28 197 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 198 218 POTENTIAL.
FT DOMAIN 219 252 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 38 88 BPTI/KUNITZ INHIBITOR 1.
FT DOMAIN 133 183 BPTI/KUNITZ INHIBITOR 2.
FT DISULFID 38 88 BY SIMILARITY.
FT DISULFID 47 71 BY SIMILARITY.
FT DISULFID 63 84 BY SIMILARITY.
FT ACT_SITE 48 49 REACTIVE BOND (BY SIMILARITY).
FT DISULFID 133 183 BY SIMILARITY.
FT DISULFID 142 166 BY SIMILARITY.
FT DISULFID 158 179 BY SIMILARITY.
FT ACT_SITE 143 144 REACTIVE BOND (BY SIMILARITY).
FT CARBOHYD 57 57 N-LINKED (GLCNAC. .).
FT CARBOHYD 94 94 N-LINKED (GLCNAC. .) (POTENTIAL).
FT VARIANT 200 200 V -> L.
FT CONFLICT 3 3 /FTID=VAR_012482.
FT CONFLICT 11 11 Q -> H (IN REF. 3).
FT CONFLICT 53 53 R -> P (IN REF. 1).
FT CONFLICT 240 240 R -> K (IN REF. 3).
FT CONFLICT 240 240 D -> H (IN REF. 3).
SQ SEQUENCE 252 AA; 28228 MW; A7D3360C0BECAB2B CRC64;

Query Match 25.6%; Score 232; DB 1; Length 252;
Best Local Similarity 30.7%; Pred. No. 3e-15;
Matches 46; Conservative 24; Mismatches 56; Indels 24; Gaps 3;

Qy 22 MHSFCAFKADGDCRAIMKREFFNIIFTRQCEFIYGGCEGNQNRFSLECKMKC--TRD 79
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 34 IHDFCLVSKVVGRCRASMPPWYNVTDGSQLFVYGGDGNNSNYLTKECLKKCATVTE 93

Qy 80 NANRIKTT-----LQEKPD-----FCFLEEDPGICRGYITRYFYNO 117
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 94 NATGDLATSRNAADSSVPSAPRRQSDSHSDMFNVEEYCTANAVTGPCRASPPRVFDV 153

Qy 118 QTKQCEKRYGCGCLGNMNFETLECKN 147
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 154 ERNSCNNFIYGGCRGNKNSYRSEACMLRC 183
```

RESULT 11

ID SPT2_MOUSE STANDARD; PRT; 252 AA.
AC Q9WU03; Q9WU04; Q9WU05;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Kunitz-type protease inhibitor 2 precursor (Hepatocyte growth factor
DE activator inhibitor type 2) (HAI-2).
GN SPINT2 OR HAI2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RC STRAIN=BALEB/C;
RX MEDLINE=99160423; PubMed=10049781;
RA Itoh H., Kataoka H., Hamasuna R., Kitamura N., Koono M.;
RT "Hepatocyte growth factor activator inhibitor type 2 lacking the first
RT Kunitz-type serine proteinase inhibitor domain is a predominant
RT product in mouse but not in human."
RL Biochem. Biophys. Res. Commun. 255:740-748(1999).
CC -!- FUNCTION: INHIBITOR OF HGF ACTIVATOR.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -!- ALTERNATIVE PRODUCTS: 3 ISOFORMS; 1 (SHOWN HERE), 2 AND 3; ARE
CC PRODUCED BY ALTERNATIVE SPLICING.
CC -!- TISSUE SPECIFICITY: ISOFORM 2 IS MORE PREDOMINANTLY EXPRESSED THAN
CC ISOFORM 1.
CC -!- DOMAIN: THIS INHIBITOR CONTAINS TWO INHIBITORY DOMAINS.
CC -!- SIMILARITY: CONTAINS 2 BPTI/KUNITZ INHIBITOR DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF099016; AAD22172.1; -
CC EMBL; AF099019; AAD22173.1; -
CC EMBL; AF099020; AAD22174.1; -
CC HSSP; P05067; ICA0.
CC MGD; MGI:1338031; Spint2.
CC InterPro; IPR002223; Kunitz_BPTI.
CC Pfam; PF00014; Kunitz_BPTI; 2.
CC PRINTS; PR00759; BASICPTASE.
CC ProDom; PD000222; Kunitz_BPTI; 2.
CC SMART; SM00131; KU; 2.
CC PROSITE; PS00280; BPTI_KUNITZ_1; 2.
CC PROSITE; PS0279; BPTI_KUNITZ_2; 2.
KW Serine protease inhibitor; Repeat; Glycoprotein; Transmembrane;
KW Signal; Alternative splicing
FT SIGNAL 1 27 POTENTIAL.
FT CHAIN 28 252 KUNITZ-TYPE PROTEASE INHIBITOR 2.
FT DOMAIN 28 197 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 198 218 POTENTIAL.
FT DOMAIN 219 252 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 38 88 BPTI/KUNITZ INHIBITOR 1.
FT DOMAIN 133 183 BPTI/KUNITZ INHIBITOR 2.
FT DISULFD 38 88 BY SIMILARITY.
FT DISULFD 38 88 BY SIMILARITY.
FT ACT_SITE 48 49 BY SIMILARITY.
FT DISULFD 63 84 REACTIVE BOND (BY SIMILARITY).
FT DISULFD 133 183 BY SIMILARITY.
FT DISULFD 142 166 BY SIMILARITY.
FT DISULFD 158 179 BY SIMILARITY.
FT ACT_SITE 143 144 REACTIVE BOND (BY SIMILARITY).
FT CARBOHYD 57 57 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 94 94 N-LINKED (GLCNAC. .) (POTENTIAL).

FT VARSPLIC 37 93 MISSING (IN ISOFORM 2 AND ISOFORM 3).
FT VARSPLIC 114 128 PRKSAEDLSAIFN -> CFVELSVAALFLFYA (IN
FT ISOFORM 3).
FT VARSPLIC 129 252 MISSING (IN ISOFORM 3).
SQ SEQUENCE 252 AA; 27914 MW; B2FF4B86924DAF8F CRC64;

Query Match 25.5%; Score 231; DB 1; Length 252;
Best Local Similarity 31.3%; Pred. No. 3.8e-15;
Matches 47; Conservative 22; Mismatches 57; Indels 24; Gaps 3;

QY 22 MHSFCFAKADDDGCRAIMKREFNIETROCEEFIYGGCEGNONRFESLECKKMC----- 76
DB 34 VHS CGVGVKCRASIPRWYNTITGSCQPFYVGGCEGNNTYQSKCELDKCAGYTE 93
QY 77 -TRDNANR-----IIRTTLQOQKPD-----FCFLEEDPGICRGYITRYFYNQ 117
DB 94 NITDDMARNRNGADSSVLSVPRKSAEDLSAEIFNVEYCVKAVTGPCRAAAPPRIWYDT 153

QY 118 QTKQCFRKYGGCGLGNNMNFETLECKNIC 147
DB 154 EKNSCIFSYGGCGRKNKNSYLSQACMQHC 183

RESULT 12
AMBP_MESAU STANDARD; PRT; 349 AA.
AC Q60559; Q60558; Q9QW86; Q9QW87;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE AMBP protein precursor [Contains: Alpha-1-microglobulin; Inter-alpha-
DE trypsin inhibitor light chain (ITI-LC) (bikunin) (HI-30); Trypsatin].
GN AMBP OR ITIL.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxID=10036;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=95110820; PubMed=7529051;
RA Ide H., Itoh H., Nawa Y.;
RT "Sequencing of cDNAs encoding alpha 1-microglobulin/bikunin of
RT Mongolian gerbil and Syrian golden hamster in comparison with man and
RT other species."
RL Biochim. Biophys. Acta 1209:286-292(1994).
RN [2]
RP SEQUENCE OF 205-348, AND SUBUNITS.
RC TISSUE=Plasma, and Urine;
RX MEDLINE=97018241; PubMed=8864857;
RA Yamamoto T., Yamamoto K., Sinohara H.;
RT "Inter-alpha-trypsin inhibitor and its related proteins in Syrian
RT hamster urine and plasma."
RL J. Biochem. 120:145-152(1996).
CC -!- FUNCTION: ALPHA-1-MICROGLOBULIN OCCURS IN MANY PHYSIOLOGICAL
CC FLUIDS INCLUDING PLASMA, URINE, AND CEREBROSPINAL FLUID. IT
CC APPEARS NOT ONLY AS A FREE MONOMER BUT ALSO IN COMPLEXES WITH IGA
CC AND ALBUMIN (BY SIMILARITY).
CC -!- FUNCTION: INTER-ALPHA-TRYPsin INHIBITOR, PRESENT IN PLASMA AND
CC URINE, INHIBITS TRYPSIN, PLASMIN, AND LYSOSOMAL GRANULOCYTIC
CC ELASTASE (BY SIMILARITY).
CC -!- FUNCTION: Trypsin is a trypsin inhibitor. It inhibits blood
CC coagulation factor Xa and trypsin about 100-fold more rapidly
CC than porcine pancreatic trypsin and chymase. It is a monomer but
CC is also found in mast cells as a complex with trypsinase (By
CC similarity).
CC -!- SUBUNIT: I-ALPHA-1 plasma protease inhibitors are assembled from
CC one or two heavy chains (H1, H2 or H3) and one light chain,
CC bikunin. Inter-alpha-inhibitor (I-ALPHA-I) is composed of H1, H2
CC and bikunin, Inter-alpha-like inhibitor (I-ALPHA-LI) of H2 and
CC bikunin, and pre-alpha-inhibitor (P-ALPHA-I) of H3 and bikunin (By
CC similarity).

CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Synthesized in the liver and secreted in plasma.
CC -!- PTM: The precursor is proteolytically processed into separately functioning proteins.
CC -!- PTM: Alpha-1-microglobulin contains a covalently linked brown-yellow chromophore (By similarity).
CC -!- PTM: Heavy chains are interlinked with bikunin via a chondroitin 4-sulfate bridge to the their C-terminal aspartate (By similarity).
CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE LIPOCALIN FAMILY.
CC -!- SIMILARITY: CONTAINS 2 BPTI/KUNITZ INHIBITOR DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D31814; BAA06601.1; .
DR HSSP; P02760; LBK.
DR InterPro; IPR002223; Kunitz_BPTI.
DR InterPro; IPR002345; Lipocalin.
DR InterPro; IPR000566; Lipocalin_cytFABP.
DR Pfam; PF00014; Kunitz_BPTI; 2.
DR Pfam; PF00061; lipocalin; 1.
DR PRINTS; PR00759; BASICPTASE.
DR PRINTS; PR00179; LIPOCALIN.
DR ProDom; PD000222; Kunitz_BPTI; 2.
DR SMART; SM00131; KU; 2.
DR PROSITE; PS00280; BPTI_KUNITZ_1; 2.
DR PROSITE; PS00279; BPTI_KUNITZ_2; 2.
DR PROSITE; PS00213; LIPOCALIN; 1.
KW Glycoprotein; Plasma; Signal; Serine protease inhibitor; Repeat; Lipocalin.
FT SIGNAL 1 19 BY SIMILARITY.
FT CHAIN 20 202 ALPHA-1-MICROGLOBULIN.
FT CHAIN 205 349 INTER-ALPHA-TRYPsin INHIBITOR LIGHT CHAIN.
FT CHAIN 282 349 TRYESTATIN.
FT DOMAIN 230 280 BPTI/KUNITZ INHIBITOR 1.
FT DOMAIN 286 336 BPTI/KUNITZ INHIBITOR 2.
FT BINDING 52 52 CHROMOPHORE (BY SIMILARITY).
FT BINDING 110 110 CHROMOPHORE (BY SIMILARITY).
FT BINDING 136 136 CHROMOPHORE (BY SIMILARITY).
FT BINDING 148 148 CHROMOPHORE (BY SIMILARITY).
FT DISULFID 90 187 BY SIMILARITY.
FT DISULFID 230 280 BY SIMILARITY.
FT DISULFID 239 263 BY SIMILARITY.
FT DISULFID 255 276 BY SIMILARITY.
FT DISULFID 286 336 BY SIMILARITY.
FT DISULFID 295 319 BY SIMILARITY.
FT DISULFID 311 332 BY SIMILARITY.
FT CARBOHYD 35 35 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 114 114 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT ACT_SITE 240 241 INHIBITORY SITE (P1) (CHYMOTRYPsin, ELASTASE) (BY SIMILARITY).
FT ACT_SITE 296 297 INHIBITORY SITE (P1) (TRYPsin) (BY SIMILARITY).
FT ACT_SITE 342 342 G -> E (IN REF. 2).
FT CONFLICT 342 342
FT SEQUENCE 349 AA; 38782 MW; 8C954584B7DE728 CRC64;
Query Match 24.6%; Score 223; DB 1; Length 349;
Best Local Similarity 32.2%; Pred. No. 3.1e-14;
Matches 47; Conservative 24; Mismatches 55; Indels 20; Gaps 3;
QY 10 ITDTELPLKLMHSECFAPKADGFCRAIMKRFENITFQCEEFYGGCGNQNFESL 69
DB 219 LVTDV---LKKEDS-CQLSYSEGCPLGMIEKYYNGASMACETPHYGGCLGNNGNFENSE 273

QY 70 ECKKMKTRDNANRIIKTTLQOEKPDFCFLEDPGICRGYITRYFYNOOTKOCERFKYGG 129
DB 274 KECLQCRIVAA-----CSLPVQGPCRAYVELWAFDAAGKCVQFSYGG 318
QY 130 CLGNMNFETLECKNICEDGPNFGQ 155
DB 319 CKGNKNGFYSEKCKEYCGVPGDGYE 344
RESULT 13
AMBP_PIG STANDARD; PRT; 337 AA.
ID AC P04366; P34954;
DT 20-MAR-1987 (Rel. 04, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE AMBP protein precursor [Contains: Alpha-1-microglobulin; Inter-alpha-trypsin inhibitor light chain (ITI-LC) (Bikunin) (HI-30) (EI-14)] (Fragment).
DE DE trypsin inhibitor light chain (ITI-LC) (Bikunin) (HI-30) (EI-14)] (Fragment).
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP MEDLINE=90353595; PubMed=1696914;
RA Gebhard W., Schreitmuller T., Vetr H., Wachter E., Hochstrasser K.;
RT "Complementary DNA and deduced amino acid sequences of porcine alpha 1-microglobulin and bikunin.";
RL FEBS Lett. 269:32-36(1990).
RN [2]
RP SEQUENCE OF 2-337 FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=91113729; PubMed=1703444;
RA Tavakkol A.;
RT "Molecular cloning of porcine alpha 1-microglobulin/HI-30 reveals developmental and tissue-specific expression of two variant messenger ribonucleic acids";
RL Biochim. Biophys. Acta 1088:47-56(1991).
RN [3]
RP SEQUENCE OF 212-334.
RX MEDLINE=85225967; PubMed=2408637;
RA Hochstrasser K., Wachter E., Albrecht G.J., Reisinger P.;
RT "Kunitz-type proteinase inhibitors derived by limited proteolysis of the inter-alpha-trypsin inhibitor, X. The amino-acid sequences of the trypsin-released inhibitors from horse and pig inter-alpha-trypsin inhibitors.";
RL Biol. Chem. Hoppe-Seyler 366:473-478(1985).
CC -!- FUNCTION: ALPHA-1-MICROGLOBULIN OCCURS IN MANY PHYSIOLOGICAL FLUIDS INCLUDING PLASMA, URINE, AND CEREBROSPINAL FLUID. IT APPEARS NOT ONLY AS A FREE MONOMER BUT ALSO IN COMPLEXES WITH IGA AND ALBUMIN.
CC -!- FUNCTION: INTER-ALPHA-TRYPsin INHIBITOR, PRESENT IN PLASMA AND URINE, INHIBITS TRYPSIN, PLASMIN, AND LYSSOMAL GRANULOCYTIC ELASTASE.
CC -!- SUBUNIT: I-ALPHA-I plasma protease inhibitors are assembled from one or two heavy chains (H1, H2 or H3) and one light chain, bikunin. Inter-alpha-inhibitor (I-ALPHA-I) is composed of H1, H2 and bikunin, inter-alpha-like inhibitor (I-ALPHA-Li) of H2 and bikunin, and pre-alpha-inhibitor (P-ALPHA-I) of H3 and bikunin (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Synthesized in the liver and secreted in plasma.
CC -!- PTM: The precursor is proteolytically processed into two separately functioning proteins.
CC -!- PTM: Alpha-1-microglobulin contains covalently linked brown-yellow chromophores (By similarity).
CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE LIPOCALIN FAMILY.
CC -!- SIMILARITY: CONTAINS 2 BPTI/KUNITZ INHIBITOR DOMAINS.
CC -----


```
DR MGD; MGI:88002; Ambp.
DR InterPro; IPR002223; Kunitz_BPTI.
DR InterPro; IPR002345; Lipocalin.
DR InterPro; IPR000566; Lipocalin_cyFABP.
DR Pfam; PF00014; Kunitz_BPTI; 2.
DR Pfam; PF00061; Lipocalin; 1.
DR PRINTS; PR00759; BASICPTASE.
DR PRINTS; PR00179; LIPOCALIN.
DR ProDom; PD000222; Kunitz_BPTI; 2.
DR SMART; SM00131; KU; 2.
DR PROSITE; PS00280; BPTI_KUNITZ_1; 2.
DR PROSITE; PS0279; BPTI_KUNITZ_2; 2.
DR PROSITE; PS00213; LIPOCALIN; 1.
KW Glycoprotein; Plasma; Signal; Serine protease inhibitor; Repeat;
KW Lipocalin.
FT SIGNAL 1 19 BY SIMILARITY.
FT CHAIN 20 202 ALPHA-1-MICROGLOBULIN.
FT CHAIN 205 349 INTER-ALPHA-TRYPsin INHIBITOR LIGHT
FT CHAIN
FT DOMAIN 230 280 BPTI/KUNITZ INHIBITOR 1.
FT DOMAIN 286 336 BPTI/KUNITZ INHIBITOR 2.
FT BINDING 52 52 CHROMOPHORE (BY SIMILARITY).
FT BINDING 110 110 CHROMOPHORE (BY SIMILARITY).
FT BINDING 136 136 CHROMOPHORE (BY SIMILARITY).
FT BINDING 148 148 CHROMOPHORE (BY SIMILARITY).
FT DISULFID 90 187 BY SIMILARITY.
FT DISULFID 230 280 BY SIMILARITY.
FT DISULFID 239 263 BY SIMILARITY.
FT DISULFID 255 276 BY SIMILARITY.
FT DISULFID 286 336 BY SIMILARITY.
FT DISULFID 295 319 BY SIMILARITY.
FT DISULFID 311 332 BY SIMILARITY.
FT CARBOHYD 33 33 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 114 114 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 233 233 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT ACT_SITE 240 241 INHIBITORY SITE (P1) (CHYMOTRYPSIN,
FT ACT_SITE 296 297 ELASTASE) (BY SIMILARITY).
FT ACT_SITE 297 INHIBITORY SITE (P1) (TRYPSIN) (BY
FT ACT_SITE 297 SIMILARITY).
FT CONFLICT 65 65 Q -> S (IN REF. 2).
FT SEQUENCE 349 AA; 39070 MW; C54D9FC7375DA80B CRC64;
Query Match 24.0%; Score 217; DB 1; Length 349;
Best Local Similarity 32.9%; Pred. No. 1.2e-13;
Matches 51; Conservative 23; Mismatches 61; Indels 20; Gaps 4;
QY 1 DSEEDDEHTITDPLPLKLMHSECAFKADGGPCRAIMKRFNFTQCEEFYGGCE 60
Db 210 ESESGTEPLITGT----LKKEDS-CQLNYSSEGPCLGMRYYNGASMACETFGGCL 264
QY 61 GNORFESLECKKMCRTDRNANRIKTTLOQEKDPDFCELEDPGICRGYITRYFYNQTK 120
Db 265 GGNFNFISEKDCLOTC-----RTAA-----CNLPVQGPCRAFIKLWAFDAAG 309
QY 121 QCERFYGGCGCLGNMNFETLECKNICEDGPNFGQ 155
Db 310 KCIQFYGGCKGNGNKFYSEKECKEYGVPGDGYE 344
RESULT 15
IATR-SHEEP
ID IATR-SHEEP STANDARD; PRU; 123 AA.
AC P13371.
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Inter-alpha-trypsin inhibitor (ITri) (GIR-14) (Inhibitory fragment of
DE ITri) (Fragment).
OS Ovis aries (Sheep), and
OC Capra hircus (Goat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
```

```
OX NCBI_TaxID=9940, 9925;
RN [1]
RP SEQUENCE.
RC SPECIES=Sheep;
RX MEDLINE=87299012; PubMed=2441725;
RA Rasp G., Hochstrasser K., Wachter E., Reisinger P.W.M.;
RT "The amino-acid sequence of the trypsin-released inhibitor from sheep
inter-alpha-trypsin inhibitor.";
RL Biol. Chem. Hoppe-Seyler 368:727-731(1987).
RN [2]
RP SEQUENCE.
RC SPECIES=C.hircus;
RX MEDLINE=90105540; PubMed=2481505;
RA Rasp G., Hochstrasser K., Gerl C., Wachter E.;
RT "Primary structure of a proteinase inhibitor released from goat serum
inter-alpha-trypsin inhibitor.";
RL Biochim. Biophys. Acta 999:335-337(1989).
CC -!- FUNCTION: THIS INHIBITORY FRAGMENT, CONTAINS TWO HOMOLOGOUS DOMAINS.
LIMITED PROTEOLYSIS WITH TRYPSIN, CONTAINS A STRONG INHIBITOR OF TRYPSIN, THE
WHEREAS THE SECOND DOMAIN IS A STRONG INHIBITOR OF TRYPSIN, THE
FIRST DOMAIN INTERACTS WEAKLY WITH PMN-GRANULOCYTTIC ELASTASE AND
NOT AT ALL WITH PANCREATIC ELASTASE.
CC -!- MISCELLANEOUS: THE AMINO ACID AT POSITION P2' (17) APPEARS TO
DETERMINE THE SPECIFICITY OF THE INHIBITION OF DOMAIN 1.
INHIBITORS WITH METHIONINE IN THIS POSITION INTERACT WEAKLY WITH
CHYMOTRYPSIN AND ELASTASE; THOSE WITH LEUCINE INTERACT STRONGLY.
CC -!- SIMILARITY: CONTAINS 2 BPTI/KUNITZ INHIBITOR DOMAINS.
PIR; A29652; A29652.
DR HSSP; P02760; IBIK.
DR InterPro; IPR002223; Kunitz_BPTI.
DR Pfam; PF00014; Kunitz_BPTI; 2.
DR ProDom; PD000222; Kunitz_BPTI; 2.
DR SMART; SM00131; KU; 2.
DR PROSITE; PS00280; BPTI_KUNITZ_1; 2.
DR PROSITE; PS0279; BPTI_KUNITZ_2; 2.
KW Plasma; Glycoprotein; Serine protease inhibitor; Repeat.
FT NON_TER 1 1
FT DOMAIN 5 55 BPTI/KUNITZ INHIBITOR 1.
FT DOMAIN 61 111 BPTI/KUNITZ INHIBITOR 2.
FT DISULFID 5 55 BY SIMILARITY.
FT DISULFID 14 38 BY SIMILARITY.
FT DISULFID 30 51 BY SIMILARITY.
FT DISULFID 61 111 BY SIMILARITY.
FT DISULFID 70 94 BY SIMILARITY.
FT DISULFID 86 107 BY SIMILARITY.
FT ACT_SITE 15 16 INHIBITORY SITE (P1) (CHYMOTRYPSIN,
FT ACT_SITE 71 72 ELASTASE).
FT ACT_SITE 71 72 INHIBITORY SITE (P1) (TRYPSIN).
FT CARBOHYD 24 24 N-LINKED (GLCNAC. . .).
FT NON_TER 123 123
FT SEQUENCE 123 AA; 13686 MW; 295038173F22D2D1 CRC64;
Query Match 23.9%; Score 216.5; DB 1; Length 123;
Best Local Similarity 37.7%; Pred. No. 4.4e-14;
Matches 46; Conservative 13; Mismatches 48; Indels 15; Gaps 2;
QY 26 CAFKADGPGCAIMKRFNFTQCEEFYGGCEGNORFESLECKKMCRTDRNANRII 85
Db 5 CQLGYSGGPGCLGMFKRYNGTSMACETFYGGCMGNGNFPSEKELQTC----- 55
QY 86 KTTIQQEKDPDFCELEDPGICRGYITRYFYNQTKQCEERFYGGCGCLGNMNFETLECKN 145
Db 56 -RTVQA-----CNLPVIRGPCRAGIELWAFDAVKGKCVRFYGGCGNGNGNFSQSECKE 109
QY 146 IC 147
Db 110 YC 111
Search completed: December 2, 2002, 10:02:53
Job time : 18.0199 secs
```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 2, 2002, 10:00:15 ; Search time 57.6716 Seconds
(without alignments)
575.215 Million cell updates/sec

```
Title: US-09-741-106-19
Perfect score: 905
Sequence: 1 DSEDEEHTIITDELPLK.....ECKNICEGDPNGQVDNYGT 161
```

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL_21:*

- ```

1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp.podent:*
12: sp.virus:*
13: sp.vertebrate:*
14: sp.unclassified:*
15: sp.vrirus:*
16: sp.archaeap:*
17: sp.bacterap:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query % |      | Length | DB     | ID     | Description |
|------------|-------|---------|------|--------|--------|--------|-------------|
|            |       | Match   |      |        |        |        |             |
| 1          | 550   | 60.8    | 287  | 6      | Q28874 | Q28874 | canis famil |
| 2          | 362   | 40.0    | 396  | 13     | Q93424 | Q93424 | cyprinus ca |
| 3          | 312.5 | 34.5    | 142  | 5      | Q8WP12 | Q8WP12 | boophilus m |
| 4          | 301.5 | 33.3    | 142  | 5      | Q8WP13 | Q8WP13 | boophilus m |
| 5          | 268.5 | 29.7    | 3198 | 5      | Q908G8 | Q908G8 | manduca sex |
| 6          | 266.5 | 29.4    | 2167 | 5      | Q76840 | Q76840 | caenorhabdi |
| 7          | 258.5 | 28.6    | 1195 | 5      | Q9N343 | Q9N343 | caenorhabdi |
| 8          | 258   | 28.5    | 1572 | 5      | Q44938 | Q44938 | haemomachus |
| 9          | 257.5 | 28.5    | 2225 | 5      | Q45881 | Q45881 | caenorhabdi |
| 10         | 254.5 | 28.1    | 3060 | 5      | Q9YAV4 | Q9YAV4 | drosophila  |
| 11         | 252   | 27.8    | 838  | 5      | Q27422 | Q27422 | caenorhabdi |
| 12         | 232   | 27.8    | 838  | 5      | Q18761 | Q18761 | caenorhabdi |
| 13         | 248   | 27.4    | 2174 | 5      | Q9GGR0 | Q9GGR0 | drosophila  |
| 14         | 244   | 27.0    | 922  | 5      | Q21418 | Q21418 | caenorhabdi |
| 15         | 240.5 | 26.6    | 1599 | 5      | Q95983 | Q95983 | caenorhabdi |
| 16         | 235   | 26.0    | 1474 | 5      | Q62504 | Q62504 | caenorhabdi |

|    |       |      |      |    |        |                    |
|----|-------|------|------|----|--------|--------------------|
| 17 | 233   | 25.7 | 988  | 5  | Q22685 | Q22685 caenorhabdi |
| 18 | 226.5 | 25.0 | 1203 | 5  | Q45916 | Q45916 caenorhabdi |
| 19 | 223   | 24.6 | 342  | 13 | P70004 | P70004 xenopus lae |
| 20 | 221   | 24.4 | 1743 | 5  | Q9XWX5 | Q9XWX5 caenorhabdi |
| 21 | 217.5 | 24.0 | 352  | 11 | Q70160 | Q70160 cavia porce |
| 22 | 217   | 24.0 | 349  | 11 | Q9DBJ9 | Q9DBJ9 mus musculu |
| 23 | 217   | 24.0 | 349  | 11 | Q925W1 | Q925W1 mus musculu |
| 24 | 213.5 | 23.6 | 507  | 11 | Q9D3K4 | Q9D3K4 mus musculu |
| 25 | 213.5 | 23.6 | 507  | 11 | Q95J04 | Q95J04 mus musculu |
| 26 | 206.5 | 22.8 | 1391 | 5  | Q19021 | Q19021 caenorhabdi |
| 27 | 203   | 22.4 | 1043 | 5  | Q17644 | Q17644 caenorhabdi |
| 28 | 199.5 | 22.0 | 151  | 4  | P78491 | P78491 homo sapien |
| 29 | 199   | 22.0 | 576  | 4  | Q8TEU8 | Q8TEU8 homo sapien |
| 30 | 197.5 | 21.8 | 165  | 5  | Q964Q0 | Q964Q0 ixodes scap |
| 31 | 190.5 | 21.0 | 183  | 5  | Q16784 | Q16784 caenorhabdi |
| 32 | 189.5 | 20.9 | 805  | 5  | Q19305 | Q19305 caenorhabdi |
| 33 | 182   | 20.1 | 984  | 5  | Q9GQ21 | Q9GQ21 calliactis  |
| 34 | 182   | 20.1 | 984  | 5  | Q9GQ11 | Q9GQ11 calliactis  |
| 35 | 179   | 19.8 | 83   | 13 | Q90W98 | Q90W98 pseudonaja  |
| 36 | 175   | 19.3 | 58   | 5  | Q9TWG0 | Q9TWG0 anemonia su |
| 37 | 174.5 | 19.3 | 548  | 4  | Q96NZ8 | Q96NZ8 homo sapien |
| 38 | 173.5 | 19.2 | 1208 | 5  | Q23456 | Q23456 caenorhabdi |
| 39 | 173   | 19.1 | 169  | 6  | Q9OX57 | Q9N0X7 bos taurus  |
| 40 | 171.5 | 19.0 | 684  | 5  | Q965M1 | Q965M1 caenorhabdi |
| 41 | 169.5 | 18.7 | 154  | 6  | Q9NOX3 | Q9N0X3 ovis aries  |
| 42 | 169.5 | 18.7 | 372  | 13 | Q91484 | Q91484 salmo salar |
| 43 | 167.5 | 18.5 | 770  | 6  | Q9TU10 | Q9TU10 sus scrofa  |
| 44 | 167   | 18.5 | 547  | 4  | Q13764 | Q13764 homo sapien |
| 45 | 167   | 18.5 | 547  | 4  | Q13764 | Q13764 homo sapien |
| 46 | 167   | 18.5 | 607  | 11 | Q99K32 | Q99K32 mus musculu |

## ALIGNMENTS

## RESULT 1

```

QZ8874 PRELIMINARY; PRT; 396 AA.
AC Q28874;
ID Q28874;
AD Q28874;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-WAR-2002 (TREMBlrel. 20, Last annotation update)
DE Tissue factor pathway inhibitor.
DE DE
DE Canis familiaris (Dog)
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_Taxid=9615;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95071310; PubMed=7980463;
RA Girard T.J., Gallani D., Broze G.J.Jr.;
RT "Complementary DNA sequencing of canine tissue factor pathway
RT inhibitor reveals a unique nanomeric repetitive sequence between the
RT second and third Kunitz domains.";
RL Biochem. J. 303:923-928(1994).
RR EMBL: S75369; AAB32443.1; -.
DR HSP; P10646; 1TFX.
DR InterPro; IPR002223; Kunitz_BPTI.
DR Prfam; PF00014; Kunitz_BPTI; 3.
DR PRINTS; PR00759; BASICPTASE.
DR ProDom; PD000222; Kunitz_BPTI; 3.
DR SMART; SM00131; KU; 3.
DR PROSITE; PS00280; BPTI_KUNITZ_1; 3.
DR PROSITE; PS02879; BPTI_KUNITZ_2; 3.
KW Serine protease inhibitor.
KW
SQ SEQUENCE 396 AA; 43948 MW; 50F65C8337A003D9 CRC64;

Query Match 60.8%; Score 550; DB 6; Length 396;
Best Local Similarity 65.8%; Pred. No. 1.3e-52;
Matches 98; Conservative 24; Mismatches 21; Indels 6; Gaps
 3
QY 3 EEDEHTIIITDELPLPKLMHSFCFAKADGPGCAIRMKRFFFTFTQCSEFTYGCGEN 62
 :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 31 DESEEYPGITD-ELPPLRIILHSCALKADNGPCRAMIRNFNFTHTQCCEFTYGGCGN 99
```





```
RESULT 7
Q9N343
ID Q9N343 PRELIMINARY; PRT; 1195 AA.
AC Q9N343;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 131.3 kDa protein.
GN Y55F3BR.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Bradshaw-Cordum H., Leonard S., Graves T.;
RT "The sequence of C. elegans cosmid Y55F3BR.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Waterston R.;
RT "Direct Submission.";
RL EMBL: AC024830; AAF59608.1; -;
DR HSSP; P00981; 1DPK.
DR InterPro; IPR002223; Kunitz_BPTI.
DR InterPro; IPR002899; WRI/EB.
DR Pfam; PF01683; EB; 1.
DR Pfam; PF00014; Kunitz_BPTI; 5.
DR PRINTS; PR00759; BASICPTASE.
DR ProDom; PD000222; Kunitz_BPTI; 4.
DR SMART; SM00131; KU; 5.
DR SMART; SM00289; WRI; 7.
DR PROSITE; PS00280; BPTI_KUNITZ_1; 1.
DR PROSITE; PS00279; BPTI_KUNITZ_2; 5.
DR PROSITE; PS00279; BPTI_KUNITZ_2; 5.
KW Hypothetical protein; Serine protease inhibitor.
SQ SEQUENCE 1195 AA; 131342 MW; E77C3A6DF2272A18 CRC64;

Query Match 28.6%; Score 258.5; DB 5; Length 1195;
Best Local Similarity 31.6%; Pred. No. 7.5e-20;
Matches 60; Conservative 26; Mismatches 45; Indels 59; Gaps 7;

QY 6 EEHTITDPELPKLMHSFCAKADGCPRAIMKRFNFITRQCEFIYGGCGNQR 65
DB 741 EKTICTE-----PLRV-----GCKOSVQRFWNAETKTCESFLITGCGGNRR 785
QY 66 FESLECKMKTRDANRIKTTLQKPDF----- 96
DB 786 FNSLNEQSYCKNINAEP--KCPQGRAYVDSGKPMQCGGLGGTACFANVECTFDGLVY 843
QY 97 -----CFLEEDPGI-C-RGYITRFYNOQTQCFERFYGGCLGNMNNFETLECKN 145
DB 844 GCCPSKAYTCSLVNKGKIGCGSSRYYYNNAKEQCSYFLGCDGNSNFPSTIEKQN 903
QY 146 ICEDG--PNG 153
DB 904 YCEIACPNG 913

RESULT 8
O44938
ID O44938 PRELIMINARY; PRT; 1572 AA.
AC O44938;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Thrombospondin.
GN THRI.
OS Haemochus contortus (Barber pole worm).
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;
OC Trichostrongyloidea; Haemonchidae; Haemonchinae; Haemonchus.
OX NCBI_TaxID=6289;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MOREDUN;
RA Skuce P.J., Newlands G.F.J., Stewart M., Pettit D., Smith D.,
RA Knox D.P.;
RT "Cloning and characterization of thrombospondin, a novel multidomain
RT glycoprotein associated with the gut of Haemonchus contortus.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF043121; AAB99830.2; -;
DR HSSP; P05067; 1CA0.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR002223; Kunitz_BPTI.
DR InterPro; IPR000884; TSPI.
DR Pfam; PF00014; Kunitz_BPTI; 6.
DR Pfam; PF00090; tsp_1; 6.
DR PRINTS; PR00759; BASICPTASE.
DR ProDom; PD000222; Kunitz_BPTI; 6.
DR SMART; SM00001; EGF_like; 1.
DR SMART; SM00131; KU; 6.
DR SMART; SM00209; TSPI; 7.
DR PROSITE; PS00280; BPTI_KUNITZ_1; 6.
DR PROSITE; PS00279; BPTI_KUNITZ_2; 6.
DR PROSITE; PS00092; TSPI; 2.
KW Serine protease inhibitor.
SQ SEQUENCE 1572 AA; 171871 MW; 2260B30DC2F903EC CRC64;

Query Match 28.5%; Score 258; DB 5; Length 1572;
Best Local Similarity 35.6%; Pred. No. 1.1e-19;
Matches 53; Conservative 22; Mismatches 58; Indels 16; Gaps 3;

QY 2 SEDEHTITDPELPKLMH--SFCAKADGCPRAIMKRFNFITRQCEFIYGGC 59
DB 1436 SRECENTCVRHSEPHSDTSHGTSVCDEAKETGPTNFATKWKYKADGTCNRFHYGGC 1495
QY 60 EGNONFESLECKMKTRDANRIKTTLQKPDFCFLEEDPGICRGYITRFYNOQT 119
DB 1496 EGTNRNEDNEQSCKAAC-----QDACTLPKVQGPCSGKHEYYENTVS 1541
QY 120 KQCFERFYGGCLGNMNNFETLECKNICE 148
DB 1542 MSCEKFTYGGCLGNTNRFSTLDECQSRQ 1570

RESULT 9
O45881
ID O45881 PRELIMINARY; PRT; 2225 AA.
AC O45881;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE W01F3.3 protein.
GN W01F3.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC Cummings P.N.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
```







RT "The sequence of C. elegans cosmid F30H5.";  
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.

RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BRISTOL N2;  
RA Waterston R.;  
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U29096; AAA68408.1; -.  
DR HSSP; P10646; IADZ.  
DR InterPro; IPR002198; ADH\_short.  
DR InterPro; IPR000561; EGF-like.  
DR InterPro; IPR002223; Kunitz\_BPTI.  
DR InterPro; IPR002221; WAP.  
DR InterPro; IPR002899; WRI/EB.  
DR Pfam; PF01683; EB; 3.  
DR Pfam; PF00014; Kunitz\_BPTI; 5.  
DR PRINTS; PR00759; BASICPTASE.  
DR ProDom; PD000222; Kunitz\_BPTI; 5.  
DR SMART; SM00181; EGF; 1.  
DR SMART; SM00131; KU; 5.  
DR SMART; SM00217; WAP; 1.  
DR SMART; SM00289; WRI; 15.  
DR PROSITE; PS00061; ADH\_SHORT; UNKNOWN\_1.  
DR PROSITE; PS00280; BPTI\_KUNITZ\_1; 2.  
DR PROSITE; PS0279; BPTI\_KUNITZ\_2; 5.  
KW Serine protease inhibitor.  
SQ SEQUENCE 1599 AA; 171659 MW; AB5E6A1D86F9880D CRC64;

Query Match 26.8%; Score 240.5; DB 5; Length 1599;  
Best Local Similarity 31.2%; Pred. No. 1e-17;  
Matches 55; Conservative 21; Mismatches 49; Indels 51; Gaps 6;

QY 18 PLKLMHSFCFAKADGPCRAIMKRFNIFTROCEEFYGGCEGNONRPFSELECKK--- 74

Db 566 PLRL-----GCKOSVRYWYNATRACEIFDYTGCGNDNNFETLECCNTCE 614

QY 75 -----MCTRDNANRIKTTLQ-----QEKPDFCFLEE 101

Db 615 NIPEPCPGDAYKDYQNVYVCSNAGSCPVNYECYFDGYVWGCCPKAYTCTLSP 674

QY 102 DPGI-C-RGITRYFYNQTKQCFERYGGCLGNMNNFETLECKNICEDEG--PNG 153

Db 675 HKGVTCGSGSYRYYNSTQECESQYNGCDGNSNNFNATREDCGYGVGCGPNG 730

Search completed: December 2, 2002, 10:04:16  
Job time : 61.6716 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 2, 2002, 09:59:15 ; Search time 10.2985 Seconds  
(without alignments)  
297.593 Million cell updates/sec

Title: US-09-741-106-10

Perfect score: 114

Sequence: 1 KTRKRKRKQKVIAEIEFVKNM 23

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_101002.\*  
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID          | Description         |
|------------|-------|-------------|--------|-------------|---------------------|
| 1          | 114   | 100.0       | 23     | 17 AAW06879 | Glycosaminoglycan   |
| 2          | 114   | 100.0       | 23     | 18 AAW21910 | Smooth muscle fibre |
| 3          | 114   | 100.0       | 37     | 17 AAR92263 | TFPI C-terminal ta  |
| 4          | 114   | 100.0       | 161    | 22 AAU02982 | Angiotensin conver  |
| 5          | 114   | 100.0       | 276    | 17 AAR92265 | TFPI mutetin, Lys36 |
| 6          | 114   | 100.0       | 276    | 17 AAR92012 | TFPI mutetin K36R.  |
| 7          | 114   | 100.0       | 276    | 18 AAW30311 | Recombinant non-g1  |
| 8          | 114   | 100.0       | 276    | 19 AAW61535 | Human tissue facto  |
| 9          | 114   | 100.0       | 277    | 14 AAR37312 | Non-glycosylated T  |
| 10         | 114   | 100.0       | 304    | 10 AAP92002 | Human tissue facto  |

|    |      |       |     |             |                     |
|----|------|-------|-----|-------------|---------------------|
| 11 | 114  | 100.0 | 304 | 16 AAR81884 | Lipoprotein-associ  |
| 12 | 114  | 100.0 | 304 | 16 AAR78389 | Human lipoprotein-  |
| 13 | 114  | 100.0 | 304 | 16 AAR67994 | Tissue factor path  |
| 14 | 114  | 100.0 | 304 | 20 AAY49557 | Human lipoprotein   |
| 15 | 114  | 100.0 | 304 | 21 AAY70272 | Human mutant tissu  |
| 16 | 114  | 100.0 | 304 | 21 AAY70273 | Human tissue facto  |
| 17 | 114  | 100.0 | 352 | 17 AAR92011 | Ubiquitin-TFPI fus  |
| 18 | 113  | 99.1  | 304 | 14 AAR42309 | LAC1 gene product.  |
| 19 | 109  | 95.6  | 23  | 18 AAW21929 | Smooth muscle fibre |
| 20 | 108  | 94.7  | 23  | 18 AAW21918 | Smooth muscle fibre |
| 21 | 107  | 93.9  | 23  | 17 AAR92266 | Truncated TFPI C-t  |
| 22 | 107  | 93.9  | 23  | 18 AAW21919 | Smooth muscle fibre |
| 23 | 103  | 90.4  | 22  | 18 AAW21928 | Smooth muscle fibre |
| 24 | 100  | 87.7  | 23  | 18 AAW21916 | Smooth muscle fibre |
| 25 | 99.5 | 87.3  | 25  | 18 AAW21912 | Smooth muscle fibre |
| 26 | 98   | 86.0  | 21  | 18 AAW21927 | Smooth muscle fibre |
| 27 | 94   | 82.5  | 20  | 18 AAW21926 | Smooth muscle fibre |
| 28 | 88   | 77.2  | 19  | 18 AAW21925 | Smooth muscle fibre |
| 29 | 84   | 73.7  | 18  | 18 AAW21924 | Smooth muscle fibre |
| 30 | 81   | 71.1  | 19  | 18 AAW21917 | Smooth muscle fibre |
| 31 | 79   | 69.3  | 16  | 18 AAW21911 | Smooth muscle fibre |
| 32 | 72   | 63.2  | 261 | 12 AAR11171 | Ser-(Aspl-Glu245)-  |
| 33 | 64   | 56.1  | 13  | 15 AAR55841 | LAC1 C-terminal (L  |
| 34 | 64   | 56.1  | 262 | 12 AAR11172 | Ser-(Aspl-Ser248)-  |
| 35 | 63   | 55.3  | 14  | 18 AAW21923 | Smooth muscle fibre |
| 36 | 63   | 55.3  | 23  | 18 AAW21915 | Smooth muscle fibre |
| 37 | 59   | 51.8  | 12  | 15 AAR45680 | Human TFPI heparin  |
| 38 | 56   | 49.1  | 12  | 18 AAW21914 | Smooth muscle fibre |
| 39 | 56   | 49.1  | 272 | 21 AAG29929 | Arabidopsis thalia  |
| 40 | 55.5 | 48.7  | 265 | 12 AAR11170 | Ser-(Aspl-Ile253)-  |
| 41 | 55   | 48.2  | 24  | 18 AAW21920 | Smooth muscle fibre |
| 42 | 54   | 47.4  | 11  | 18 AAW21913 | Smooth muscle fibre |
| 43 | 54   | 47.4  | 12  | 18 AAW21922 | Smooth muscle fibre |
| 44 | 52.5 | 46.1  | 14  | 18 AAW21906 | Smooth muscle fibre |
| 45 | 50   | 43.9  | 95  | 22 AAW72157 | Human bone marrow   |

#### ALIGNMENTS

RESULT 1

AAW06879

ID AAW06879 standard; Peptide; 23 AA.

XX AAW06879;

XX 18-MAR-1997 (first entry)

XX Glycosaminoglycan binding peptide from TFPI.

XX Complement inhibitor; membrane co-factor protein; MCP;

XX decay accelerating factor; DAF; chimeric protein; glycosaminoglycan;

XX heparin; cell lysis; sepsis; adult respiratory distress syndrome;

XX reperfusion injury; cell damage; tissue factor pathway inhibitor;

XX TFPI.

XX Synthetic.

XX WO9634965-A2.

XX 07-NOV-1996.

XX 03-MAY-1996; 95WO-US06301.

XX 05-MAY-1995; 95US-0435149.

XX (CHIR ) CHIRON CORP.

XX Creasey AA, Innis MA, Zaror I;

XX WPI; 1996-506167/50.

XX Chimeric proteins for inhibiting complement-mediated cell lysis

PT

PT comprise membrane co-factor protein and decay accelerating factor  
XX peptide sequences

PS Claim 11; Page 26; 33pp; English.

XX A glycosaminoglycan binding peptide (AAW06879) derived from tissue  
XX factor pathway inhibitor is used in novel chimeric proteins of  
XX formula A-R1-B-R2-C, where A and C are peptides (AAW06875-79,  
XX AAW06883-90) capable of binding glycosaminoglycans (esp. heparin)  
XX present on cell surfaces, R1 is membrane co-factor protein (MCP) or  
XX decay accelerating factor (DAF), R2 is DAF when R1 is MCP or MCP  
XX when R1 is DAF, and B is a peptide that may have complement  
XX inhibitor activity. The chimeric proteins (see also AAW06882) are  
XX directed to cell surfaces where they inhibit complement-mediated  
XX cell lysis. They are used to treat and prevent disease states in  
XX which complement plays a role, e.g. sepsis, adult respiratory  
XX distress syndrome, reperfusion injury and tissue damage.

XX Sequence 23 AA;

Query Match 100.0%; Score 114; DB 17; Length 23;  
Best Local Similarity 100.0%; Pred. No. 3.7e-10;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTKRRKKQKVKIAYEEIFVKNM 23  
XXXXXXXXXXXXXXXXXXXX  
DB 1 KTKRRKKQKVKIAYEEIFVKNM 23

RESULT 2

AAW21910

ID AAW21910 standard; peptide; 23 AA.

XX AC AAW21910;

XX DT 14-JAN-1998 (first entry)

XX Smooth muscle fibre proliferation inhibitor peptide 1.

XX Generic; TFPI; inhibitor; proliferation; smooth muscle cell; prevention;  
XX treatment; arteriosclerosis; restenosis; angioplasty; luminal stenosis;  
XX vascular transplantation; leiomyosarcoma; human.

XX Synthetic.

XX WO9715598-A1.

XX PD 01-MAY-1997.

XX PF 23-OCT-1996; 96WO-JP03080.

XX PR 24-OCT-1995; 95JP-0300792.

XX PA (KAGA ) CEMO-SERO-THERAPEUTIC RES INST.

XX PI Hara S, Kamikubo Y, Miyamoto S, Nakahara Y, Takemoto S;

XX DR WPI; 1997-258960/23.

XX Peptide(s) comprising basic rich peptide bound at C-terminal to  
XX consecutive hydrophobic rich peptide - useful for inhibiting smooth  
XX muscle fibre cell proliferation

XX Claim 18; Page 22; 45pp; Japanese.

XX Novel peptides contain: (a) a peptide sequence (P1) rich in basic amino  
XX acid residues (preferably lysine, arginine or histidine); and (b) a  
XX peptide sequence (P2) containing at least two consecutive hydrophobic  
XX amino acid residues (preferably phenylalanine, isoleucine, leucine,  
XX methionine, proline, valine, tryptophan or tyrosine); where P2 is bound  
XX to the C-terminal end of P1 either directly or through a linker sequence  
XX of several amino acid residues. The peptides inhibit proliferation of  
XX smooth muscle cells, and are useful in the prevention and treatment of

CC arteriosclerosis associated with smooth muscle cell proliferation,  
CC restenosis after angioplasty, luminal stenosis after vascular  
CC transplantation, and leiomyosarcoma. AAW21910-12 are specific inhibitory  
CC peptides based on the generic P1 and P2 sequences shown in AAW21905 and  
XX AAW21909.

XX Sequence 23 AA;

Query Match 100.0%; Score 114; DB 18; Length 23;  
Best Local Similarity 100.0%; Pred. No. 3.7e-10;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTKRRKKQKVKIAYEEIFVKNM 23

XXXXXXXXXXXXXXXXXXXX

DB 1 KTKRRKKQKVKIAYEEIFVKNM 23

RESULT 3

AAW22263

ID AAR92263 standard; peptide; 37 AA.

XX AC AAR92263;

XX DT 30-SEP-1996 (first entry)

XX DE TFPI C-terminal tail.

XX Kunitz-type domain; C-terminal tail; tissue factor pathway inhibitor;  
XX TFPI; TFPI-2; cell surface localisation; glycosaminoglycan; heparin;  
XX phospholipid; binding; chimeric protein; mutein; substitution;  
XX P1-reactive site; sepsis; septic shock; thrombosis; up-regulation;  
XX tissue factor; injury; trauma; endotoxin; TNF; cancer; IL-1;  
XX tumour necrosis factor; interleukin.

XX Homo sapiens.

XX WO9604378-A2.

XX PD 15-FEB-1996.

XX PF 25-JUL-1995; 95WO-US09464.

XX PR 05-AUG-1994; 94US-0286521.

XX PA (CHIR ) CHIRON CORP.

XX PI Creasey AA, Innis MA;

XX WPI; 1996-129394/13.

XX Chimeric protein comprising Kunitz-type domains from TFPI-1 and -2  
XX - used for the treatment of septic shock and thrombosis disorders

XX Claim 10; Page 6; 68pp; English.

XX The sequences given in AAR92257-64 represent first, second and third  
XX Kunitz-type domains and the C-terminal tails derived from tissue  
XX factor pathway inhibitor (TFPI) and TFPI-2. These sequences are  
XX highly basic and may be involved in cell surface localisation by  
XX glycosaminoglycan (including heparin) or phospholipid binding.  
XX These sequences may be used in the construction of the chimeric  
XX proteins or muteins of the invention. These muteins have one or  
XX more substitutions exclusively in the P1-reactive site of one or  
XX more Kunitz-type domains. The chimeric proteins may comprise an  
XX alternative glycosaminoglycan binding peptide selected from those given  
XX in AAR92266-73. The chimeric proteins and muteins may be used in a  
XX pharmaceutical composition for the treatment of sepsis, septic shock and  
XX thrombosis disorders. The proteins may be generally useful in the  
XX treatment of diseases caused by the up-regulation of tissue factor  
XX brought on by injury, trauma, endotoxin, TNF, cancer, IL-1 or other  
XX agents or conditions.

XX Sequence 37 AA;





CC purifying and refolding proteins (especially TFPI) which have been  
CC engineered by genetic recombination and produced in bacterial, yeast or  
CC other cells in a form that has a non-native tertiary structure. TFPI is a  
CC coagulation inhibitor which has clot-inhibiting properties.

XX Sequence 276 AA;

Query Match 100.0%; Score 114; DB 18; Length 276;  
Best Local Similarity 100.0%; Pred. No. 4.8e-09;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTKRRKKQKRVKIAYEEIFVKNM 23  
|||||  
DB 254 KTKRRKKQKRVKIAYEEIFVKNM 276

RESULT 8  
AAW61535  
ID AAW61535 standard; protein; 276 AA.

XX AAW61535;

XX 06-NOV-1998 (first entry)

XX Human tissue factor pathway inhibitor (TFPI).

XX Human tissue factor pathway inhibitor; TFPI; TFPI-2; cell proliferation;  
XX angiogenesis-related disease; cancer; arthritis; macular degeneration;  
XX diabetic retinopathy.

OS Homo sapiens.

FH Key Location/Qualifiers

FT Modified-site 2  
FT /note= "Potential phosphorylation site"

FT Domain 26..76

FT /note= "Kunitz-1 type protease inhibitor domain"

FT Disulfide-bond 26..76

FT Disulfide-bond 35..59

FT Disulfide-bond 51..72

FT Domain 97..147

FT /note= "Kunitz-2 type protease inhibitor domain"

FT Disulfide-bond 97..147

FT Disulfide-bond 106..130

FT Modified-site 117

FT /note= "N-glycosylated"

FT Disulfide-bond 122..143

FT Modified-site 167

FT /note= "N-glycosylated"

FT Domain 189..239

FT /note= "Kunitz-3 type protease inhibitor domain"

FT Disulfide-bond 189..239

FT Disulfide-bond 198..222

FT Disulfide-bond 214..235

FT Modified-site 228

FT /note= "N-glycosylated"

XX W09834634-A1.

XX 13-AUG-1998.

XX 06-FEB-1998; 98WO-US02699.

XX 06-FEB-1997; 97US-0796850.

XX (ENTR-) ENTREMED INC.

XX Green SJ, Papathanassiou AE;

XX WPI; 1998-446947/38.

XX Composition comprising tissue factor pathway inhibitor for

XX inhibiting cell proliferation - for treating angiogenesis related

PT diseases e.g. cancer, arthritis, macular degeneration and diabetic  
PT retinopathy

XX Claim 6; Pages 23-24; 37pp; English.

XX The present sequence represents the human tissue factor pathway  
CC inhibitor (TFPI). The invention provides compositions using TFPI  
CC and its homologs, e.g. TFPI-2 (AAW61536), for inhibiting cell  
CC proliferation. The compositions are claimed to be useful for inhibiting  
CC an angiogenesis-related disease, such as cancer, arthritis, macular  
CC degeneration or diabetic retinopathy.

XX Sequence 276 AA;

Query Match 100.0%; Score 114; DB 19; Length 276;  
Best Local Similarity 100.0%; Pred. No. 4.8e-09;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTKRRKKQKRVKIAYEEIFVKNM 23  
|||||  
DB 254 KTKRRKKQKRVKIAYEEIFVKNM 276

RESULT 9

AAAR37312

ID AAR37312 standard; protein; 277 AA.

XX AAR37312;

XX 20-SEP-1993 (first entry)

XX Non-glycosylated TFPI.

XX Non-glycosylated; tissue factor pathway inhibitor; TFPI; multivalent;  
XX lipoprotein associated coagulation inhibitor; inhibitor; Kunitz-type;  
XX coagulation; domain; factor VIIa; LACI; tissue factor; factor Xa;  
XX complex; coagulation.

XX Escherichia coli.

FH Key Location/Qualifiers

FT Disulfide-bond 26..76

FT Disulfide-bond 51..72

FT Disulfide-bond 97..147

FT Disulfide-bond 106..130

FT Disulfide-bond 122..143

FT Disulfide-bond 189..239

FT Disulfide-bond 198..222

FT Disulfide-bond 214..235

XX US5212091-A.

XX 18-MAY-1993.

XX 02-MAR-1992; 92US-0844297.

XX 02-MAR-1992; 92US-0844297.

XX (MONS ) MONSANTO CO.

XX Diaz-Collier JA, Gustafson ME, Wun T;

XX WPI; 1993-175458/21.

XX Prodn. of non-glycosylated form of tissue factor pathway

XX inhibitor in high yield - comprises culturing E.coli cells

XX transformed with replication expression vector and subjecting

XX isolated inclusion bodies to sulphatolysis or redn. with

XX beta-mercapto-ethanol, etc.

XX Claim 1; Column 15-18; 25pp; English.

XX This sequence represents a non-glycosylated form of tissue factor

CC pathway inhibitor (TFPI). TFPI is alternatively known as lipo-  
 CC protein associated coagulation inhibitor (LACI). TFPI is a multi-  
 CC valent kunitz-type inhibitor of coagulation. The primary amino acid  
 CC sequence of TFPI shows that it contains a highly negatively charged  
 CC amino terminus, three tandem kunitz-type inhibitory domains and a  
 CC highly positively charged carboxyl terminal. The first kunitz domain  
 CC of TFPI is needed for the inhibition of factor VIIa/tissue factor  
 CC complex and the second kunitz domain of TFPI is responsible for the  
 CC inhibition of factor Xa. The function of the third kunitz domain  
 CC is unknown. TFPI is thought to act in vivo to limit the initiation  
 CC of coagulation by forming an inert, quaternary factor Xa:TFPI:factor  
 CC VIIa:tissue factor complex.

XX Sequence 277 AA;  
 SQ Query Match 100.0%; Score 114; DB 14; Length 277;  
 Best Local Similarity 100.0%; Pred. No. 4.9e-09;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTKRRKKQKRVKIAVEEIFVKNM 23  
 |||||  
 DB 255 KTKRRKKQKRVKIAVEEIFVKNM 277

RESULT 10  
 AAP92002  
 ID AAP92002 standard; protein; 304 AA.  
 XX  
 AC AAP92002;  
 XX  
 DT 09-FEB-1990 (first entry)  
 XX  
 DE Human tissue factor inhibitor (TFI).  
 DE Human tissue factor inhibitor; TFI; human placenta lambda-P9 clone; basic  
 KW protease inhibitor gene superfamily  
 KW  
 XX Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Region 1..28  
 FT /label= Signal\_region  
 FT /note= "A-T rich"  
 FT Cleavage-site 28..29  
 FT /note= "Possible site for signal peptidase"  
 FT Protein 29..304  
 FT Region 145..147  
 FT /note= "Potential N-linked glycosylation site"  
 FT Region 195..197  
 FT /note= "Potential N-linked glycosylation site"  
 FT Region 256..258  
 FT /note= "Potential N-linked glycosylation site"  
 FT Region 31..53  
 FT /note= "See comments below"  
 FT Region 82..88  
 FT /note= "See comments below"  
 FT Region 153..16  
 FT /note= "See comments below"  
 XX EP318451-A.  
 XX  
 PN 31-MAY-1989.  
 XX  
 PD 22-JUL-1988; 88EP-0870127.  
 XX  
 PF 23-NOV-1987; 87US-0123753.  
 XX  
 PR (MONS ) MONSANTO CO (UNIW).  
 XX  
 PI Broze GJ, Kretzmer KK, Wun TC;  
 XX WPI; 1989-159483/22.  
 DR N-PSDB; AAN90108.

XX DNA encoding human tissue factor inhibitor  
 PT - used in study of coagulation cascade for agents  
 PT which inhibit factor Xa and Factor VIIa-TF  
 XX  
 PS Claim 3; Figure 3; 14pp; English.  
 XX  
 CC Amino acid sequence of tissue factor inhibitor (TFI) cDNA isolate from  
 CC lambda-P9 clone of human placenta cDNA library. Domains discerned  
 CC include: highly negatively charged N-terminal; highly positively charged  
 CC carboxy-terminal; intervening portion consisting of 3 homologous domains  
 CC with sequences typical of kunitz-type enzyme inhibitors. Based on  
 CC homology study, it appears to be a member of the basic protease  
 CC inhibitor gene superfamily. Sequences in misc. regions in feature table  
 CC above have been independently confirmed by amino acid sequence analysis.

XX Sequence 304 AA;  
 SQ Query Match 100.0%; Score 114; DB 10; Length 304;  
 Best Local Similarity 100.0%; Pred. No. 5.4e-09;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTKRRKKQKRVKIAVEEIFVKNM 23  
 |||||  
 DB 282 KTKRRKKQKRVKIAVEEIFVKNM 304

RESULT 11  
 AAR81884  
 ID AAR81884 standard; protein; 304 AA.  
 XX  
 AC AAR81884;  
 XX  
 DT 18-MAR-1996 (first entry)  
 XX  
 DE Lipoprotein-associated coagulation inhibitor (LACI).  
 KW Lipoprotein-associated coagulation inhibitor; LACI; kallikrein;  
 KW inhibitor; KIP; Kunitz domain; hereditary angioedema.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 1..28  
 FT /label= sig\_peptide  
 FT Domain 50..107  
 FT /note= "Kunitz domain LACI-K1"  
 FT Domain 121..178  
 FT /note= "Kunitz domain LACI-K2"  
 FT Domain 213..270  
 FT /note= "Kunitz domain LACI-K3"  
 XX WO9521601-A2.  
 PN  
 XX 17-AUG-1995.  
 PD  
 XX 11-JAN-1995; 95WO-US00299.  
 PF  
 XX 10-MAR-1994; 94US-0208264.  
 PR  
 XX 11-JAN-1994; 94US-0179964.  
 XX  
 PA (PROT-) PROTEIN ENG CORP.  
 XX  
 PI Ladner RC, Markland W;  
 XX  
 DR WPI; 1995-292934/38.  
 XX  
 PT Kallikrein inhibiting proteins comprising a kunitz domain homologous  
 PT to bovine pancreatic trypsin inhibitor - useful for preventing or  
 PT treating disorders attributable to excessive kallikrein activity,  
 PT eg. in hereditary angioedema.  
 XX  
 PS Disclosure; Page 24; 46pp; English.

XX AAR81884 is the human lipoprotein-associated coagulation inhibitor  
 CC LACI. The Kunitz domain, LACI-K1, of LACI is a kallikrein  
 CC inhibiting protein (KIP) upon which the claimed KIPs of the invention  
 CC are based. The KIPs can be used for treating or preventing disorders  
 CC attributable to excessive kallikrein activity, e.g. hereditary  
 CC angioedema. The KIPs can also be used for assaying, purifying and in  
 CC vivo imaging of kallikrein.

XX Sequence 304 AA;  
 CC  
 CC Query Match 100.0%; Score 114; DB 16; Length 304;  
 CC Best Local Similarity 100.0%; Pred. No. 5.4e-09;  
 CC Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTKRRKKQKRVKIAYEEIFVKNM 23  
 DB 282 KTKRRKKQKRVKIAYEEIFVKNM 304

RESULT 12  
 AAR78389  
 ID AAR78389 standard; protein; 304 AA.  
 XX  
 AC AAR78389;  
 XX  
 DT 27-FEB-1996 (first entry)  
 XX  
 DE Human lipoprotein-associated coagulation inhibitor.  
 XX  
 KW Human; lipoprotein-associated coagulation inhibitor; peptide library;  
 KW inhibitor; plasmin; bovine; pancreatic trypsin inhibitor; Kunitz domain;  
 KW fibrinolysis; fibrinogenolysis; bleeding; thrombolytic.  
 XX  
 OS Homo sapiens.

XX Key Location/Qualifiers  
 FH Peptide 1..28  
 FT /note= "signal peptide"  
 FT Domain 50..107  
 FT /note= "LACI-K1 domain"  
 FT Domain 121..178  
 FT /note= "LACI-K2 domain"  
 FT Domain 213..270  
 FT /note= "LACI-K3 domain"

XX WO9518830-A2.  
 XX  
 PD 13-JUL-1995.  
 XX  
 PF 11-JAN-1995; 95WO-US00298.  
 XX  
 PR 10-MAR-1994; 94US-0208265.  
 PR 11-JAN-1994; 94US-0179658.  
 XX  
 PA (PROT-) PROTEIN ENG CORP.  
 PI Ladner RC, Markland W;  
 XX  
 DR WPI; 1995-255042/33.  
 XX  
 PT Novel plasmin inhibiting protein comprising a Kunitz Domain - useful  
 PT to prevent/treat disorders attributable to excess plasmin activity.  
 XX  
 PS Disclosure; Page 28; 59pp; English.

XX The amino acid sequence of the human lipoprotein-associated coagulation  
 CC inhibitor (LACI). The sequence encoding the K1 domain was used to  
 CC generate a library of peptides which act as inhibitors of plasmin. The  
 CC peptides (see AAR78390-R78599) are mutants homologous to bovine  
 CC pancreatic trypsin inhibitor (BPTI) Kunitz domains that inhibit plasmin.  
 CC The peptides can be used to prevent or treat a clinical condition  
 CC exacerbated by plasmin e.g. inappropriate fibrinolysis or

CC fibrinogenolysis, excessive bleeding associated with thrombolytics.  
 XX  
 SQ Sequence 304 AA;  
 CC  
 CC Query Match 100.0%; Score 114; DB 16; Length 304;  
 CC Best Local Similarity 100.0%; Pred. No. 5.4e-09;  
 CC Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTKRRKKQKRVKIAYEEIFVKNM 23  
 DB 282 KTKRRKKQKRVKIAYEEIFVKNM 304

RESULT 13  
 AAR67994  
 ID AAR67994 standard; Protein; 304 AA.  
 XX  
 AC AAR67994;  
 XX  
 DT 13-AUG-1995 (first entry)  
 XX  
 DE Tissue factor pathway inhibitor.  
 XX  
 KW Yeast aspartic protease 3; YAP3; signal peptide; protein secretion;  
 KW tissue factor pathway inhibitor; TFPI.  
 XX  
 OS Homo sapiens.

XX Key Location/Qualifiers  
 FH Peptide 1..28  
 FT /label= Sig\_peptide  
 FT /note= "TFPI signal peptide"  
 XX  
 PN WO9502059-A.  
 XX  
 PD 19-JAN-1995.  
 XX  
 PF 08-JUL-1994; 94WO-DK00281.  
 XX  
 PR 08-JUL-1993; 93DK-0000828.  
 XX  
 PA (NOVO ) NOVO-NORDISK AS.  
 XX  
 PI Christiansen L, Petersen JG;  
 XX  
 DR WPI; 1995-066903/09.  
 DR N-PSDB; AAQ81396.  
 XX  
 PT DNA construct encoding the yeast aspartic protease 3 signal  
 PT peptide - provides improved secretion of proteins in transformed  
 PT yeast cells, such as aprotinin and insulin  
 XX  
 PS Disclosure; Page 27-28; 56pp; English.  
 XX  
 CC A Sali fragment encoding human tissue factor pathway inhibitor  
 CC (TFPI) is given in AAQ81396. It was used to construct expression  
 CC vectors allowing production of TFPI in Saccharomyces cerevisiae as  
 CC fusions to the yeast aspartic protease 3 (YAP3) signal peptide,  
 CC facilitating secretion of the recombinant TFPI.  
 XX  
 SQ Sequence 304 AA;  
 CC  
 CC Query Match 100.0%; Score 114; DB 16; Length 304;  
 CC Best Local Similarity 100.0%; Pred. No. 5.4e-09;  
 CC Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTKRRKKQKRVKIAYEEIFVKNM 23  
 DB 282 KTKRRKKQKRVKIAYEEIFVKNM 304

RESULT 14  
 AAY49557

ID AAY49557 standard; Protein; 304 AA.  
 AC AAY49557;  
 XX  
 DT 13-JAN-2000 (first entry)  
 XX  
 DE Human lipoprotein associated coagulation inhibitor protein sequence.  
 XX  
 KW Human; coding sequence polymorphism; vascular pathology gene;  
 KW polymorphic site; phenotype correlation; forensic; paternity testing;  
 KW medicine; genetic analysis; vascular disease.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W0950454-A2.  
 XX  
 PD 07-OCT-1999.  
 XX  
 PF 26-MAR-1999; 99WO-US06473.  
 XX  
 PR 01-APR-1998; 98US-0054272.  
 XX  
 PA (WHED ) WHITEHEAD INST BIOMEDICAL RES.  
 XX  
 PI Lander ES, Daley GO, Cargill M, Ireland JS, Rozen SG;  
 XX  
 DR WPI: 1999-620066/53.  
 DR N-PSDB; AAZ32166.  
 XX  
 PT Determination of polymorphisms in genes, especially those identifying  
 PT predisposition to vascular disease -  
 XX  
 PS Disclosure; Fig 10; 134pp; English.  
 XX  
 CC AA32159 to AA32194 represent reference alleles for specifically  
 CC claimed nucleic acid sequences from the present invention which comprise  
 CC polymorphic sites as given in a table in the specification, selected  
 CC from 92 single nucleotide polymorphisms in which the nucleotide at the  
 CC polymorphic site is different from a nucleotide at the same site in a  
 CC reference allele. The nucleic acids, and primers and probes, are used to  
 CC identify polymorphisms, which may predispose an individual to disease,  
 CC especially a vascular disease. They can also be used in phenotype  
 CC correlations, forensics, paternity testing, medicine or genetic  
 CC analysis. AAY49550 to AAY49573 represent the proteins which correspond  
 CC to some of the reference alleles.  
 XX  
 SQ Sequence 304 AA;  
 Query Match 100.0%; Score 114; DB 20; Length 304;  
 Best Local Similarity 100.0%; Pred. No. 5.4e-09;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KTKRRKKQKVIAEYIFVKNM 23  
 |||||  
 DB 282 KTKRRKKQKVIAEYIFVKNM 304  
 RESULT 15  
 AAY70272  
 ID AAY70272 standard; Protein; 304 AA.  
 AC AAY70272;  
 XX  
 DT 06-JUN-2000 (first entry)  
 XX  
 DE Human mutant tissue factor pathway inhibitor protein (TFPI).  
 XX  
 KW Tissue factor pathway inhibitor; TFPI; mutant; human; regulator;  
 KW Kunitz-type proteinase inhibitor; extrinsic blood coagulation pathway;  
 KW diagnosis; thrombotic disorder; thromboembolic disease.  
 XX  
 OS Homo sapiens.  
 XX

FH Location/Qualifiers  
 FT 1..28  
 FT Peptide  
 FT /label= Signal\_peptide  
 FT 29..304  
 FT Protein  
 FT /label= Mature\_human\_mutant\_TFPI\_protein  
 FT /note= "Tissue factor pathway inhibitor"  
 FT 29..53  
 FT Region  
 FT /label= Acidic\_amino\_terminal\_region  
 FT /note= "Consists of negatively charged amino acids"  
 FT 54..104  
 FT Domain  
 FT /label= Kunitz\_type\_proteinase\_inhibitory\_domain-1  
 FT /note= "Binding site for factor VIIa/tissue factor"  
 FT 54..104  
 FT Disulfide-bond  
 FT 63..87  
 FT Disulfide-bond  
 FT 79..100  
 FT Disulfide-bond  
 FT 105..124  
 FT Region  
 FT /note= "Connecting chain-1"  
 FT 125..175  
 FT Domain  
 FT /label= Kunitz\_type\_proteinase\_inhibitory\_domain-2  
 FT /note= "Binds activated factor Xa"  
 FT 125..175  
 FT Disulfide-bond  
 FT 134..158  
 FT Disulfide-bond  
 FT 145  
 FT Modified-site  
 FT /note= "N-glycosylation site"  
 FT 150..171  
 FT Disulfide-bond  
 FT 176..216  
 FT Region  
 FT /note= "Connecting chain-2"  
 FT 179  
 FT Misc-difference  
 FT /note= "Wild type Pro substituted with Leu"  
 FT 195  
 FT Modified-site  
 FT /note= "N-glycosylation site"  
 FT 202  
 FT Modified-site  
 FT /note= "O-glycosylation site"  
 FT 203  
 FT Modified-site  
 FT /note= "O-glycosylation site"  
 FT 217..267  
 FT Domain  
 FT /label= Kunitz\_type\_proteinase\_inhibitory\_domain-3  
 FT 217..267  
 FT Disulfide-bond  
 FT 226..250  
 FT Disulfide-bond  
 FT 242..263  
 FT Disulfide-bond  
 FT 268..304  
 FT Region  
 FT /label= Basic\_carboxy\_terminal\_region  
 FT /note= "Consists of positively charged amino acids"  
 XX  
 XX WO200011034-A1.  
 PN  
 XX  
 XX 02-MAR-2000.  
 PD  
 XX  
 XX 18-AUG-1999; 99WO-EP06054.  
 PF  
 XX  
 XX 25-AUG-1998; 98EP-0115957.  
 PR  
 XX  
 XX (KLEE/) KLEESTEK K.  
 PA  
 XX  
 XX Kleesiek K, Brinkmann T, Prohaska W, Goetting C, Schmidt M;  
 PI  
 XX  
 XX WPI: 2000-237617/20.  
 DR  
 XX  
 XX N-PSDB; AA251316.  
 DR  
 XX  
 XX Polynucleotide encoding tissue factor pathway inhibitor mutant useful  
 PT for diagnosing a disposition for venous thromboembolic diseases in  
 PT humans -  
 PT  
 XX  
 XX Claim 10; Page 22-23; 27pp; English.  
 PS  
 XX  
 CC The present amino acid sequence is the mutant human tissue factor pathway  
 CC inhibitor (TFPI) protein. A single nucleotide substitution (C-T) in exon  
 CC 7 of the TFPI wild type gene, leads to a proline-leucine substitution in  
 CC the mature peptide. TFPI is a single chain glycoprotein present in plasma  
 CC in trace amounts, that belongs to the class of Kunitz-type proteinase  
 CC inhibitors. It is an important regulator in the extrinsic blood  
 CC coagulation pathway. The mutant TFPI sequence has a new recognition site  
 CC for the restriction enzyme BseNI, that is useful for in vitro diagnosis



CC of thrombotic disorders, especially thromboembolic diseases, by screening  
 CC genomic DNA from blood samples.

XX

SQ Sequence 304 AA;

Query Match 100.0%; Score 114; DB 21; Length 304;  
 Best Local Similarity 100.0%; Pred. No. 5.4e-09;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTKRKKKORVKIAYEEIFVKNN 23

Db 282 KTKRKKKORVKIAYEEIFVKNN 304

Search completed: December 2, 2002, 10:02:23  
 Job time : 11.2985 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 2, 2002, 10:02:30 ; Search time 3.77612 Seconds  
(without alignments)  
179.212 Million cell updates/sec

Title: US-09-741-106-10  
Perfect score: 114  
Sequence: 1 KTKRRKKQKVIAEYEFVKNM 23

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description       |
|------------|-------|-------------|--------|----|-------------------|
| 1          | 114   | 100.0       | 23     | 1  | US-08-437-841-10  |
| 2          | 114   | 100.0       | 23     | 1  | US-08-286-521-10  |
| 3          | 114   | 100.0       | 23     | 1  | US-08-436-175-10  |
| 4          | 114   | 100.0       | 23     | 2  | US-08-435-149-15  |
| 5          | 114   | 100.0       | 23     | 4  | US-08-943-682-10  |
| 6          | 114   | 100.0       | 23     | 4  | US-09-051-986-1   |
| 7          | 114   | 100.0       | 23     | 5  | PCT-US95-09464-10 |
| 8          | 114   | 100.0       | 24     | 4  | US-09-051-986-20  |
| 9          | 114   | 100.0       | 37     | 1  | US-08-437-841-7   |
| 10         | 114   | 100.0       | 37     | 1  | US-08-286-521-7   |
| 11         | 114   | 100.0       | 37     | 1  | US-08-436-175-7   |
| 12         | 114   | 100.0       | 37     | 4  | US-08-943-682-7   |
| 13         | 114   | 100.0       | 37     | 5  | PCT-US95-09464-7  |
| 14         | 114   | 100.0       | 276    | 1  | US-07-828-920A-1  |
| 15         | 114   | 100.0       | 276    | 1  | US-08-437-841-9   |
| 16         | 114   | 100.0       | 276    | 1  | US-08-286-521-9   |
| 17         | 114   | 100.0       | 276    | 1  | US-08-436-175-9   |
| 18         | 114   | 100.0       | 276    | 2  | US-08-796-850-1   |
| 19         | 114   | 100.0       | 276    | 3  | US-08-854-764-3   |
| 20         | 114   | 100.0       | 276    | 4  | US-08-943-682-9   |
| 21         | 114   | 100.0       | 276    | 5  | PCT-US95-09377-3  |
| 22         | 114   | 100.0       | 276    | 5  | PCT-US95-09464-9  |
| 23         | 114   | 100.0       | 277    | 1  | US-07-844-297-1   |
| 24         | 114   | 100.0       | 304    | 1  | US-08-026-145-2   |
| 25         | 114   | 100.0       | 304    | 1  | US-08-446-646-9   |
| 26         | 114   | 100.0       | 304    | 1  | US-08-676-125A-18 |
| 27         | 114   | 100.0       | 304    | 2  | US-09-136-012A-18 |

|    |     |       |     |   |                   |                    |
|----|-----|-------|-----|---|-------------------|--------------------|
| 28 | 114 | 100.0 | 304 | 3 | US-08-676-124-1   | Sequence 1, Appl   |
| 29 | 114 | 100.0 | 304 | 3 | US-08-208-264A-25 | Sequence 25, Appl  |
| 30 | 114 | 100.0 | 304 | 3 | US-09-414-878-1   | Sequence 1, Appl   |
| 31 | 114 | 100.0 | 304 | 3 | US-09-240-136-1   | Sequence 1, Appl   |
| 32 | 114 | 100.0 | 304 | 4 | US-09-054-782-2   | Sequence 2, Appl   |
| 33 | 114 | 100.0 | 304 | 4 | US-09-421-097-25  | Sequence 25, Appl  |
| 34 | 114 | 100.0 | 304 | 4 | US-09-638-770A-1  | Sequence 1, Appl   |
| 35 | 114 | 100.0 | 304 | 6 | 5466783-2         | Patent No. 5466783 |
| 36 | 114 | 100.0 | 352 | 3 | US-08-854-764-2   | Sequence 2, Appl   |
| 37 | 114 | 100.0 | 352 | 5 | PCT-US95-09377-2  | Sequence 2, Appl   |
| 38 | 110 | 96.5  | 23  | 4 | US-09-051-986-5   | Sequence 5, Appl   |
| 39 | 109 | 95.6  | 23  | 4 | US-09-051-986-19  | Sequence 19, Appl  |
| 40 | 108 | 94.7  | 23  | 4 | US-09-051-986-8   | Sequence 8, Appl   |
| 41 | 107 | 93.9  | 23  | 4 | US-09-051-986-9   | Sequence 9, Appl   |
| 42 | 103 | 90.4  | 22  | 4 | US-09-051-986-18  | Sequence 18, Appl  |
| 43 | 100 | 87.7  | 23  | 4 | US-09-051-986-6   | Sequence 6, Appl   |
| 44 | 98  | 86.0  | 21  | 4 | US-09-051-986-17  | Sequence 17, Appl  |
| 45 | 94  | 82.5  | 20  | 4 | US-09-051-986-16  | Sequence 16, Appl  |

ALIGNMENTS

RESULT 1  
US-08-437-841-10  
; Sequence 10, Application US/08437841  
; Patent No. 5563123  
; GENERAL INFORMATION:  
; APPLICANT: Innis, Michael  
; APPLICANT: Creasey, Abia  
; TITLE OF INVENTION: Chimeric Proteins  
; NUMBER OF SEQUENCES: 37  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Chiron Corporation  
; STREET: 4560 Horton St.  
; CITY: Emeryville  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94608  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30B  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/437,841  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/286,521  
; FILING DATE: 05-AUG-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Saveriede, Paul B.  
; REGISTRATION NUMBER: 36,914  
; REFERENCE/DOCKET NUMBER: 0990.001  
; TELEPHONE: 510-601-2585  
; TELEFAX: 510-655-3542  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 23 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-437-841-10

Query Match 100.0%; Score 114; DB 1; Length 23;  
Best Local Similarity 100.0%; Pred. No. 1.9e-10;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTKRRKKQKVIAEYEFVKNM 23  
|||||

Db 1 KTKRRKKQKVKIAYEEIFVKNM 23

US-08-286-521-10  
US-08-286-521-10, Application US/08286521  
Patent No. 5589359  
GENERAL INFORMATION:  
APPLICANT: Innis, Michael  
APPLICANT: Creasey, Abia  
TITLE OF INVENTION: Chimeric Proteins  
NUMBER OF SEQUENCES: 37  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Chiron Corporation  
STREET: 4560 Horton St.  
CITY: Emeryville  
STATE: CA  
COUNTRY: USA  
ZIP: 94608  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30B  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/286,521  
FILING DATE: 05-AUG-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Savereide, Paul B.  
REGISTRATION NUMBER: 36,914  
REFERENCE/DOCKET NUMBER: 0990.001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 510-601-2585  
TELEFAX: 510-655-3542  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 23 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-436-175-10

Query Match 100.0%; Score 114; DB 1; Length 23;  
Best Local Similarity 100.0%; Pred. No. 1.9e-10;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTKRRKKQKVKIAYEEIFVKNM 23  
Db 1 KTKRRKKQKVKIAYEEIFVKNM 23

RESULT 4  
US-08-435-149-15  
Sequence 15, Application US/08435149  
Patent No. 5866402  
GENERAL INFORMATION:  
APPLICANT: INNIS, MICHAEL A.  
APPLICANT: ZAROR, ISABEL  
APPLICANT: CREASEY, ABIA A.  
TITLE OF INVENTION: CHIMERIC MCP AND DAF PROTEINS WITH CELL  
SURFACE LOCALIZING DOMAIN  
NUMBER OF SEQUENCES: 26  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CHIRON CORPORATION  
STREET: INTELLECTUAL PROPERTY - R440, P.O. BOX 8097  
CITY: EMERYVILLE  
STATE: CALIFORNIA  
COUNTRY: U.S.A.  
ZIP: 94662-8097  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/435,149  
FILING DATE: 05-MAY-1995  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: SAVEREIDE, PAUL B.  
REGISTRATION NUMBER: 36,914  
REFERENCE/DOCKET NUMBER: 0989.001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (510) 601-2585  
TELEFAX: (510) 655-3542  
TELEX: N/A  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 23 amino acids  
TYPE: amino acid  
STRANDEDNESS: single

QY 1 KTKRRKKQKVKIAYEEIFVKNM 23  
Db 1 KTKRRKKQKVKIAYEEIFVKNM 23

RESULT 3  
US-08-436-175-10  
Sequence 10, Application US/08436175  
Patent No. 5696088  
GENERAL INFORMATION:  
APPLICANT: Innis, Michael  
APPLICANT: Creasey, Abia  
TITLE OF INVENTION: Chimeric Proteins  
NUMBER OF SEQUENCES: 37  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Chiron Corporation  
STREET: 4560 Horton St.  
CITY: Emeryville  
STATE: CA  
COUNTRY: USA  
ZIP: 94608  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-435-149-15

Query Match 100.0%; Score 114; DB 2; Length 23;  
Best Local Similarity 100.0%; Pred. No. 1.9e-10;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTKRRKKQVKIAYEEIFVKNM 23

Db 1 KTKRRKKQVKIAYEEIFVKNM 23

## RESULT 5

US-08-943-682-10  
; Sequence 10, Application US/08943682  
; Patent No. 6174721

## ; GENERAL INFORMATION:

; APPLICANT: Innis, Michael  
; APPLICANT: Creasey, Abba  
; TITLE OF INVENTION: Chimeric Proteins  
; NUMBER OF SEQUENCES: 37  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Chiron Corporation  
; STREET: 4560 Horton St.  
; CITY: Emeryville  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94608  
; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30B  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/943,682  
; FILING DATE: 03-OCT-1997  
; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/438,184  
; FILING DATE: 09-MAY-1995  
; APPLICATION NUMBER: US 08/286,521  
; FILING DATE: 05-AUG-1994  
; ATTORNEY/AGENT INFORMATION:

; NAME: Savereide, Paul B.  
; REGISTRATION NUMBER: 36,914  
; REFERENCE/DOCKET NUMBER: 0990.001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 510-601-2585  
; TELEFAX: 510-655-3542

; INFORMATION FOR SEQ ID NO: 10:

; SEQUENCE CHARACTERISTICS:  
; LENGTH: 23 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-943-682-10

Query Match 100.0%; Score 114; DB 4; Length 23;  
Best Local Similarity 100.0%; Pred. No. 1.9e-10;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTKRRKKQVKIAYEEIFVKNM 23

Db 1 KTKRRKKQVKIAYEEIFVKNM 23

## RESULT 6

US-09-051-986-1  
; Sequence 1, Application US/09051986A  
; Patent No. 6191113

## ; GENERAL INFORMATION:

; APPLICANT: NAKAHARA, Yo  
; APPLICANT: HARA, Saburo  
; APPLICANT: KAMIKUBO, Yuichi  
; APPLICANT: TAKEMOTO, Sumiyo  
; APPLICANT: MIYAMOTO, Seiji  
; TITLE OF INVENTION: NOVEL PEPTIDE  
; FILE REFERENCE: NAKAHARA-1  
; CURRENT APPLICATION NUMBER: US/09/051,986A  
; CURRENT FILING DATE: 1998-04-24  
; EARLIER APPLICATION NUMBER: JP 300792/1995  
; EARLIER FILING DATE: 1995-10-24  
; EARLIER APPLICATION NUMBER: PCT/JP96/03080  
; EARLIER FILING DATE: 1996-10-23  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 23  
; TYPE: PRT  
; ORGANISM: Human Tissue Factor Pathway Inhibitor  
US-09-051-986-1

Query Match 100.0%; Score 114; DB 4; Length 23;  
Best Local Similarity 100.0%; Pred. No. 1.9e-10;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTKRRKKQVKIAYEEIFVKNM 23

Db 1 KTKRRKKQVKIAYEEIFVKNM 23

## RESULT 7

PCT-US95-09464-10  
; Sequence 10, Application PC/TUS9509464  
; GENERAL INFORMATION:

; APPLICANT: CHIRON CORPORATION  
; TITLE OF INVENTION: Chimeric Proteins  
; NUMBER OF SEQUENCES: 37  
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Chiron Corporation  
; STREET: 4560 Horton St.  
; CITY: Emeryville  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94608  
; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30B  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/09464  
; FILING DATE: 25-JULY-1995  
; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:  
; NAME: Savereide, Paul B.  
; REGISTRATION NUMBER: 36,914  
; REFERENCE/DOCKET NUMBER: 0990.100  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 510-601-2585  
; TELEFAX: 510-655-3542

; INFORMATION FOR SEQ ID NO: 10:

; SEQUENCE CHARACTERISTICS:  
; LENGTH: 23 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
PCT-US95-09464-10

Query Match 100.0%; Score 114; DB 5; Length 23;  
Best Local Similarity 100.0%; Pred. No. 1.9e-10;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 11  
US-08-436-175-7  
: Sequence 7, Application US/08436175



Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTKRKKQKRVKIAEYEEIFVKNM 23  
 |||||  
 Db 15 KTKRKKKQKRVKIAEYEEIFVKNM 37

Db 15 KTKRKKKORVKIAYEEIFVKNM 37

## RESULT 14

US-07-828-920A-1  
Sequence 1, Application US/07828920A  
Patent No. 5312736  
GENERAL INFORMATION:  
APPLICANT: Rasmussen, Jesper  
APPLICANT: No. 5312736dfang, Ole Juul  
TITLE OF INVENTION: Anticoagulant Protein  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: No. 53127360 No. 5312736disk of No. 5312736th America, Inc.  
STREET: 405 Lexington Avenue, Suite 6200  
CITY: New York  
STATE: New York  
COUNTRY: United States of America  
ZIP: 10174-6201

Query Match 100.0%; Score 114; DB 1: Length 276;

Best Local Similarity 100.0%; Pred. No. 2.3e-09;

|         |     |              |    |            |    |        |    |      |    |
|---------|-----|--------------|----|------------|----|--------|----|------|----|
| Matches | 23; | Conservative | 0; | Mismatches | 0; | Indels | 0; | Gaps | 0; |
|---------|-----|--------------|----|------------|----|--------|----|------|----|

QY 1 KTKRKKQVRVKIAYEEIFVKNM 23  
|||||  
db 254 KTKRKKQVRVKIAYEEIFVKNM 276

Db 254 KTKRRKKORVKIAYEEIFVKNM 276

## RESULT 15

US-08-437-841-9  
Sequence 9, Application US/08437841  
Patent No. 5563123  
GENERAL INFORMATION:  
APPLICANT: Innis, Michael

```

; APPLICANT: Creasey, Abba
; TITLE OF INVENTION: Chimeric Proteins
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton St.
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/437,841
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/286,521
; FILING DATE: 05-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Savereide, Paul B.
; REGISTRATION NUMBER: 36,914
; REFERENCE/DOCKET NUMBER: 0990.001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 510-601-2585
; TELEFAX: 510-655-3542
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 276 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; PS-08-437-841-9

```

Query Match 100.0%; Score 114; DB 1; Length 276;

Best Local Similarity 100.0%; Pred. No. 2.3e-09;

|         |    |              |   |            |   |        |   |      |   |
|---------|----|--------------|---|------------|---|--------|---|------|---|
| Matches | 23 | Conservative | 0 | Mismatches | 0 | Indels | 0 | Gaps | 0 |
|---------|----|--------------|---|------------|---|--------|---|------|---|

QY 1 KTKRKRKKQRVKIAEIEIFVKNM 23  
|||||  
Db 254 KTKRKRKKQRVKIAEIEIFVKNM 276

db 254 KTKRRKKORVKIAYEEIFVKNM 276

Search completed: December 2, 2002, 10:05:57  
Job time : 4.77612 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 2, 2002, 10:00:50 : Search time 2.17413 Seconds  
(without alignments)  
168.461 Million cell updates/sec

Title: US-09-741-106-10

Perfect score: 114

Sequence: 1 KTKRRKKQKVIAVEEIVKNN 23

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 102317 seqs, 15924203 residues

Total number of hits satisfying chosen parameters: 102317

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PublishedApplications\_AA:\*  
1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score | Query Match % | Length | ID | Description         |
|------------|-------|---------------|--------|----|---------------------|
| 1          | 114   | 100.0         | 23     | 9  | US-10-086-176A-3    |
| 2          | 114   | 100.0         | 30     | 9  | US-10-086-176A-2    |
| 3          | 114   | 100.0         | 45     | 9  | US-10-086-176A-1    |
| 4          | 114   | 100.0         | 276    | 9  | US-10-086-176A-5    |
| 5          | 114   | 100.0         | 276    | 10 | US-09-766-778-1     |
| 6          | 50    | 43.9          | 95     | 10 | US-09-864-761-47012 |
| 7          | 42.5  | 37.3          | 201    | 10 | US-09-925-297-714   |
| 8          | 42.5  | 37.3          | 357    | 10 | US-09-907-537-2     |
| 9          | 42    | 36.8          | 76     | 9  | US-10-002-344A-229  |
| 10         | 42    | 36.8          | 999    | 9  | US-09-895-913A-226  |
| 11         | 41    | 36.0          | 56     | 10 | US-09-864-761-41214 |
| 12         | 41    | 36.0          | 57     | 10 | US-09-864-761-41286 |
| 13         | 41    | 36.0          | 63     | 10 | US-09-764-846-177   |
| 14         | 41    | 36.0          | 89     | 10 | US-09-864-761-35241 |
| 15         | 41    | 36.0          | 101    | 10 | US-09-864-761-35241 |
| 16         | 41    | 36.0          | 129    | 10 | US-09-864-761-33599 |
| 17         | 41    | 36.0          | 238    | 10 | US-09-925-301-1411  |
| 18         | 41    | 36.0          | 242    | 10 | US-09-815-242-5463  |
| 19         | 41    | 36.0          | 499    | 10 | US-09-815-242-12164 |
|            |       |               |        |    | Sequence 35385, A   |

|    |    |      |      |    |                     |                    |
|----|----|------|------|----|---------------------|--------------------|
| 20 | 40 | 35.1 | 25   | 10 | US-09-864-761-47439 | Sequence 47439, A  |
| 21 | 40 | 35.1 | 47   | 10 | US-09-864-761-40474 | Sequence 40474, A  |
| 22 | 40 | 35.1 | 52   | 10 | US-09-864-761-46109 | Sequence 46109, A  |
| 23 | 40 | 35.1 | 54   | 10 | US-09-864-761-45531 | Sequence 45531, A  |
| 24 | 40 | 35.1 | 67   | 12 | US-10-001-843-156   | Sequence 156, App  |
| 25 | 40 | 35.1 | 71   | 10 | US-09-864-761-37061 | Sequence 37061, A  |
| 26 | 40 | 35.1 | 85   | 10 | US-09-864-761-36403 | Sequence 36403, A  |
| 27 | 40 | 35.1 | 86   | 10 | US-09-864-761-33832 | Sequence 33832, A  |
| 28 | 40 | 35.1 | 130  | 10 | US-09-864-761-34671 | Sequence 34671, A  |
| 29 | 40 | 35.1 | 131  | 10 | US-09-864-761-35536 | Sequence 35536, A  |
| 30 | 40 | 35.1 | 141  | 10 | US-09-864-761-36893 | Sequence 36893, A  |
| 31 | 40 | 35.1 | 216  | 10 | US-09-925-295-841   | Sequence 841, App  |
| 32 | 40 | 35.1 | 241  | 10 | US-09-938-803-8     | Sequence 8, Appli  |
| 33 | 40 | 35.1 | 361  | 12 | US-10-015-498-2     | Sequence 2, Appli  |
| 34 | 40 | 35.1 | 545  | 10 | US-09-978-242-1     | Sequence 1, Appli  |
| 35 | 40 | 35.1 | 561  | 10 | US-09-821-687-4     | Sequence 4, Appli  |
| 36 | 40 | 35.1 | 579  | 10 | US-09-925-300-1415  | Sequence 1415, Ap  |
| 37 | 40 | 35.1 | 911  | 10 | US-09-828-423-4     | Sequence 4, Appli  |
| 38 | 40 | 35.1 | 1005 | 10 | US-09-925-301-1335  | Sequence 1335, Ap  |
| 39 | 40 | 35.1 | 1407 | 10 | US-09-815-242-10439 | Sequence 10439, A  |
| 40 | 40 | 35.1 | 1407 | 10 | US-09-815-242-14042 | Sequence 14042, A  |
| 41 | 39 | 34.2 | 37   | 10 | US-09-925-300-1099  | Sequence 1099, Ap  |
| 42 | 39 | 34.2 | 72   | 10 | US-09-864-761-35548 | Sequence 35548, A  |
| 43 | 39 | 34.2 | 171  | 10 | US-09-812-133-2     | Sequence 2, Appli  |
| 44 | 39 | 34.2 | 190  | 9  | US-09-813-398-8     | Sequence 8, Appli  |
| 45 | 39 | 34.2 | 214  | 10 | US-09-349-954A-22   | Sequence 22, Appli |

## ALIGNMENTS

RESULT 1  
US-10-086-176A-3  
; Sequence 3, Application US/10086176A  
; Patent No. US20020173465A1  
; GENERAL INFORMATION:  
; APPLICANT: Hembrough, Todd  
; APPLICANT: Papathanassiou, Adonia E.  
; APPLICANT: Greep, Shawn J.  
; TITLE OF INVENTION: Compositions and Methods for Inhibiting Cellular Proliferation  
; FILE REFERENCE: 05213-0296 43170-266780  
; CURRENT APPLICATION NUMBER: US/10/086,176A  
; CURRENT FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: US 09/766,778  
; PRIOR FILING DATE: 2001-01-22  
; PRIOR APPLICATION NUMBER: US 09/227,955  
; PRIOR FILING DATE: 1999-01-11  
; PRIOR APPLICATION NUMBER: US 08/796,850  
; PRIOR FILING DATE: 1997-02-06  
; PRIOR APPLICATION NUMBER: US 09/130,273  
; PRIOR FILING DATE: 1998-08-06  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 3  
; LENGTH: 23  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic peptide  
US-10-086-176A-3

Query Match 100.0% Score 114: DB 9: Length 23;  
Best Local Similarity 100.0%; Pred. No. 9.1e-10;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTKRRKKQKVIAVEEIVKNN 23  
|||||

Db 1 KTKRRKKQKVIAVEEIVKNN 23  
|||||

RESULT 2  
US-10-086-176A-2



FILE FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Greene, Jamie L.  
REGISTRATION NUMBER: 32,467  
REFERENCE/DOCKET NUMBER: 05213-0290  
TELEPHONE: (404) 818-3700  
TELEFAX: (404) 818-3799  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 276 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: Active-site  
LOCATION: 2..3  
OTHER INFORMATION: /note= "Site of partial phosphorylation"  
FEATURE:  
NAME/KEY: Active-site  
LOCATION: 117..118  
OTHER INFORMATION: /note= "Potential site for N-linked glycosylation"  
FEATURE:  
NAME/KEY: Active-site  
LOCATION: 167..168  
OTHER INFORMATION: /note= "Potential site for N-linked glycosylation"  
FEATURE:  
NAME/KEY: Active-site  
LOCATION: 228..229  
OTHER INFORMATION: /note= "Potential site for N-linked glycosylation"  
FEATURE:  
NAME/KEY: Domain  
LOCATION: 26..76  
OTHER INFORMATION: /label= Kunitz-1  
FEATURE:  
NAME/KEY: Domain  
LOCATION: 97..147  
OTHER INFORMATION: /label= Kunitz-2  
FEATURE:  
NAME/KEY: Domain  
LOCATION: 189..239  
OTHER INFORMATION: /label= Kunitz-3  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-766-778-1

Query Match 100.0%; Score 114; DB 10; Length 276;  
Best Local Similarity 100.0%; Pred. No. 1.2e-08;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KTKRKRKRQKRVKIAVEIFVKNM 23  
|||  
Db 254 KTKRKRKRQKRVKIAVEIFVKNM 276

RESULT 6  
US-09-864-761-47012  
; Sequence 47012, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO  
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
; FILE REFERENCE: Aeomica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 47012  
; LENGTH: 95  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AC005600.1  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.54  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.45  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.41  
; OTHER INFORMATION: EST\_HUMAN HIT: AW853983.1, EVALUE 2.00e-51  
; OTHER INFORMATION: SWISSPROT HIT: P96095, EVALUE 1.90e-01  
US-09-864-761-47012

Query Match 43.9%; Score 50; DB 10; Length 95;  
Best Local Similarity 64.3%; Pred. No. 1.5;  
Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 3 KKKRKRQKRVKIAVE 16  
|||  
Db 25 KPRKRAQLRVAYE 38

RESULT 7  
US-09-925-297-714  
; Sequence 714, Application US/09925297  
; Patent No. US20020081659A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA105

; CURRENT APPLICATION NUMBER: US/09/925,297  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05989  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 928  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 714  
; LENGTH: 201  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-925-297-714

Query Match 37.3%; Score 42.5; DB 10; Length 201;  
Best Local Similarity 43.5%; Pred. No. 33;  
Matches 10; Conservative 4; Mismatches 2; Indels 7; Gaps 1;

QY 1 KTKRKRK-----KORVKIAYE 16  
Db 47 KTKRKRRELSEEQKEIKDAFE 69

## RESULT 8

US-09-907-537-2  
; Sequence 2, Application US/09907537  
; Patent No. US20020068348A1  
; GENERAL INFORMATION:  
; APPLICANT: Meyers, Rachel  
; APPLICANT: Tsai, Fong-Ying  
; TITLE OF INVENTION: 43755, A NOVEL HUMAN METHYLTRANSFERASE  
; TITLE OF INVENTION: FAMILY MEMBER AND USES THEREOF  
; FILE REFERENCE: MNI-176  
; CURRENT APPLICATION NUMBER: US/09/907,537  
; CURRENT FILING DATE: 2001-07-16  
; PRIOR APPLICATION NUMBER: 60/218,470  
; PRIOR FILING DATE: 2000-07-14  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 357  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-907-537-2

Query Match 37.3%; Score 42.5; DB 10; Length 357;  
Best Local Similarity 50.0%; Pred. No. 60;  
Matches 11; Conservative 2; Mismatches 6; Indels 3; Gaps 1;

QY 1 KTKRKRKRQKRVKIAYEEIFVKN 22  
Db 275 KTKRKRQKRVKIAYEEIFVGN 293

## RESULT 9

US-10-002-344A-229  
; Sequence 229, Application US/10002344A  
; Patent No. US20020172959A1  
; GENERAL INFORMATION:  
; APPLICANT: Recipon, Herve  
; APPLICANT: Sun, Yongming  
; APPLICANT: Chen, Sei-Yu  
; APPLICANT: Liu, Chenghua  
; APPLICANT: Turner, Leah  
; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Proteins  
; FILE REFERENCE: DEX-0241  
; CURRENT APPLICATION NUMBER: US/10/002,344A  
; CURRENT FILING DATE: 2002-03-15  
; PRIOR APPLICATION NUMBER: US 60/242,998  
; PRIOR FILING DATE: 2000-10-25  
; NUMBER OF SEQ ID NOS: 277  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 229

; LENGTH: 76  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-10-002-344A-229

Query Match 36.8%; Score 42; DB 9; Length 76;  
Best Local Similarity 75.0%; Pred. No. 14;  
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KTKRKRKRQKRVK 12  
Db 22 KTKRKRKRKRK 33

## RESULT 10

US-09-895-913A-226  
; Sequence 226, Application US/09895913A  
; Patent No. US20020160456A1  
; GENERAL INFORMATION:  
; APPLICANT: Kleanthous, Harold  
; APPLICANT: Al-Garawi, Amal  
; APPLICANT: Miller, Charles  
; APPLICANT: Tomb, Jean Francois  
; APPLICANT: Comen, Raymond P.  
; TITLE OF INVENTION: Identification of Polynucleotides  
; TITLE OF INVENTION: Encoding No. US20020160456A1 Helicobacter Polypeptides in t  
; TITLE OF INVENTION: Genome  
; FILE REFERENCE: 06132/043002  
; CURRENT APPLICATION NUMBER: US/09/895,913A  
; CURRENT FILING DATE: 2001-06-29  
; PRIOR APPLICATION NUMBER: US 08/881,227  
; PRIOR FILING DATE: 1997-06-24  
; NUMBER OF SEQ ID NOS: 368  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 226  
; LENGTH: 999  
; TYPE: PRT  
; ORGANISM: Helicobacter pylori  
US-09-895-913A-226

Query Match 36.8%; Score 42; DB 9; Length 999;  
Best Local Similarity 43.8%; Pred. No. 2e+02;  
Matches 7; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 3 KTKRKRKRQKRVKIAYEEI 18  
Db 346 ERKKRKRKSKLNL 361

## RESULT 11

US-09-864-761-41214  
; Sequence 41214, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aecm1ca-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27

PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 41214  
LENGTH: 56  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AC009311.2  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.8  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.3  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.6  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.3  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1  
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.4  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2  
US-09-864-761-41214

Query Match 36.0%; Score 41; DB 10; Length 56;  
Best Local Similarity 52.9%; Pred. No. 14;  
Matches 9; Conservative 3; Mismatches 5; Indels 5; Gaps 0;

QY 1 KTKRRKKKORVKIAYEE 17  
: |||||: ||

DB 15 RKKKKKKKKKKKKEE 31

## RESULT 12

US-09-864-761-41286  
Sequence 41286, Application US/09864761  
Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharron G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
FILE REFERENCE: Aecmica-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263.6

PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 41286  
LENGTH: 57  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AL157381.7  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.4  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.5  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.4  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.2  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.1  
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.3  
US-09-864-761-41286

Query Match 36.0%; Score 41; DB 10; Length 57;  
Best Local Similarity 52.9%; Pred. No. 14;  
Matches 9; Conservative 3; Mismatches 5; Indels 5; Gaps 0;

QY 1 KTKRRKKKORVKIAYEE 17  
: |||||: ||

DB 1 KRKKKKKKKKKKKEE 17

## RESULT 13

US-09-764-846-177  
Sequence 177, Application US/09764846  
Patent No. US20020102638A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: PT212  
CURRENT APPLICATION NUMBER: US/09/764,846  
CURRENT FILING DATE: 2001-01-17  
Prior application data removed - consult PALM or file wrapper  
NUMBER OF SEQ ID NOS: 348  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 177  
LENGTH: 63  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-764-846-177

```
Query Match 36.0%; Score 41; DB 10; Length 63;
Best Local Similarity 46.7%; Pred. No. 16;
Matches 7; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KTKRKRKKQKVKIAY 15
 | | | | | | | | | |
Db 48 KKKKKKKKKKKKISW 62

RESULT 14
US-09-864-761-35241
; Sequence 35241, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aemica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 35241
; LENGTH: 89
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL034548.24
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.1
```

```
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.7
US-09-864-761-35241

Query Match 36.0%; Score 41; DB 10; Length 89;
Best Local Similarity 52.9%; Pred. No. 23;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 KTKRKRKKQKVKIAYEE 17
 | | | | | | | | | |
Db 28 KRKRKRKKKKKKKKKEE 44

RESULT 15
US-09-864-761-33599
; Sequence 33599, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aemica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 33599
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
```

```

; OTHER INFORMATION: MAP TO AP000161.1
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.7
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2
US-09-864-761-33599

```

```

Query Match 36.0%; Score 41; DB 10; Length 101;
Best Local Similarity 52.9%; Pred. No. 26;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 1 KTKRKRKKORVKIAYEE 17
 | :|:|:| | |
DB 3 KKKRKRKKRKKKEEEE 19

```

```

Search completed: December 2, 2002, 10:05:17
Job time : 2.17413 secs

```





GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model  
Run on: December 2, 2002, 10:00:35 ; Search time 4.1194 Seconds  
(without alignments)  
536.751 Million cell updates/sec

Title: US-09-741-106-10  
Perfect score: 114  
Sequence: 1 KTRRRKQKRVKIAYEEIFVKNM 23  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues  
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries  
Database : PIR\_73.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID       | Description         |
|------------|-------|-------------|--------|----------|---------------------|
| 1          | 114   | 100.0       | 304    | 1 TIHUGK | tissue factor path  |
| 2          | 113   | 99.1        | 304    | 1 JC2264 | tissue factor path  |
| 3          | 91    | 79.8        | 396    | 2 S53325 | tissue factor path  |
| 4          | 77    | 67.5        | 299    | 2 I45937 | tissue factor path  |
| 5          | 77    | 67.5        | 300    | 2 S12143 | lipoprotein-associ  |
| 6          | 56    | 49.1        | 272    | 2 T01480 | hypothetical prote  |
| 7          | 52    | 45.6        | 4961   | 2 T18489 | hypothetical prote  |
| 8          | 49    | 43.0        | 114    | 2 C97205 | uncharacterized pr  |
| 9          | 49    | 43.0        | 298    | 2 G90529 | heat shock protein  |
| 10         | 49    | 43.0        | 357    | 2 G69393 | conserved hypothet  |
| 11         | 48.5  | 42.5        | 135    | 2 T38956 | very hypothetical   |
| 12         | 48    | 42.1        | 231    | 2 T48215 | translation initia  |
| 13         | 48    | 42.1        | 727    | 2 T26096 | hypothetical prote  |
| 14         | 48    | 42.1        | 2025   | 2 JC5020 | tetratricopeptide   |
| 15         | 47.5  | 41.7        | 182    | 2 AC0545 | hypothetical prote  |
| 16         | 47    | 41.2        | 780    | 2 G72405 | ATP-dependent DNA   |
| 17         | 47    | 41.2        | 1113   | 2 T20004 | hypothetical prote  |
| 18         | 47    | 41.2        | 1633   | 2 JC5056 | polybromo 1 - chic  |
| 19         | 46.5  | 40.8        | 310    | 2 A33489 | hypothetical prote  |
| 20         | 46.5  | 40.8        | 402    | 2 T45518 | hypothetical prote  |
| 21         | 46.5  | 40.8        | 402    | 2 AC2304 | transposase all398  |
| 22         | 46.5  | 40.8        | 402    | 2 AC2552 | transposase all801  |
| 23         | 46.5  | 40.8        | 402    | 2 AG2506 | transposase alr722  |
| 24         | 46.5  | 40.8        | 402    | 2 AG2506 | transposase alr722  |
| 25         | 46.5  | 40.8        | 402    | 2 AI2318 | transposase alr410  |
| 26         | 46.5  | 40.8        | 413    | 2 AG2456 | transposase alr520  |
| 27         | 46.5  | 40.8        | 653    | 2 F70383 | organic solvent to  |
| 28         | 46    | 40.4        | 240    | 2 H72294 | ribonuclease III -  |
| 29         | 45    | 39.5        | 402    | 2 S47329 | OXAL1 protein precu |

ALIGNMENTS

RESULT 1

TIHUGK  
tissue factor pathway inhibitor precursor [validated] - human  
N:Alternate names: extrinsic pathway inhibitor; lipoprotein-associated coagulation in  
C:Species: Homo sapiens (man)  
C>Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 08-Dec-2000  
C:Accession: A23712; A39176; A28650; A60433; B60433; S13034; A34315; A38294; S03903  
R:Girard, T.J.; Eddy, R.; Wesselschmidt, R.L.; MacPhail, L.A.; Likert, K.M.; Byers, M  
J. Biol. Chem. 266, 5036-5041, 1991  
A>Title: Structure of the human lipoprotein-associated coagulation inhibitor gene. In  
A:Reference number: A23712; MUID:91161593; PMID:2002045  
A:Accession: A23712  
A:Molecule type: DNA  
A:Residues: 1-304 <GIR>  
A:Cross-references: GB:M59493; GB:M59499; NID:g187204; PIDN:AAA59526.1; PID:g187206  
R:van der Logt, C.P.E.; Reitsma, P.H.; Bertina, R.M.  
Biochemistry 30, 1571-1577, 1991  
A>Title: Intron-exon organization of the human gene coding for the lipoprotein-associ  
A:Reference number: A39176; MUID:91129227; PMID:1993173  
A:Accession: A39176  
A:Molecule type: DNA  
A:Residues: 1-304 <VAN>  
A:Cross-references: GB:M58650; GB:J05312; NID:g186827; PIDN:AAA59480.1; PID:g186829  
R:Wun, T.C.; Kretzmer, K.K.; Girard, T.J.; Milelich, J.P.; Broze Jr., G.J.  
J. Biol. Chem. 263, 6001-6004, 1988  
A>Title: Cloning and characterization of a cDNA coding for the lipoprotein-associated  
A:Reference number: A28650; MUID:88198127; PMID:2452157  
A:Accession: A28650  
A:Molecule type: mRNA  
A:Residues: 1-304 <WUN>  
A:Cross-references: GB:J03225; NID:g180545; PIDN:AAA50222.1; PID:g180546  
R:Girard, T.J.; Warren, L.A.; Novotny, W.F.; Bejcek, B.E.; Milelich, J.P.; Broze Jr.,  
Thromb. Res. 55, 37-50, 1989  
A>Title: Identification of the 1.4 KB and 4.0 KB messages for the lipoprotein associa  
A:Reference number: A60433; MUID:89388722; PMID:2781520  
A:Accession: A60433  
A>Status: not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-304 <G12>  
A:Experimental source: endothelial cells  
A:Accession: B60433  
A:Molecule type: protein  
A:Residues: XX, 31-53, X', 55-56 <G13>  
A:Experimental source: recombinant material from mouse Cl37 cells  
R:Girard, T.J.; McCourt, D.; Novotny, W.F.; MacPhail, L.A.; Likert, K.M.; Broze Jr.,  
Biochem. J. 270, 621-625, 1990  
A>Title: Endogenous phosphorylation of the lipoprotein-associated coagulation inhibit  
A:Reference number: S13034; MUID:91054349; PMID:2122883  
A:Accession: S13034  
A:Molecule type: protein  
A:Residues: 29-35 <G14>

30 45 39.5 532 2 S78176 transcription init  
31 45 39.5 2210 1 RRXPLC genome polypeptide  
32 44.5 39.0 667 2 S48285 probable glycine-t  
33 44 38.6 126 2 A97110 hypothetical prote  
34 44 38.6 176 2 G83851 RNA polymerase sig  
35 44 38.6 344 2 G97257 spore coat protein  
36 44 38.6 368 2 JC6081 proximal sequence  
37 44 38.6 413 2 G86181 hypothetical prote  
38 44 38.6 492 2 C59097 hypothetical prote  
39 44 38.6 1166 2 H71609 hypothetical prote  
40 44 38.6 2057 2 F90109 splicing factor pr  
41 43.5 38.2 407 2 AE2497 transposase all715  
42 43.5 38.2 471 1 T04935 translation initia  
43 43.5 38.2 517 2 T49173 hypothetical prote  
44 43 37.7 317 2 C83652 hypothetical prote  
45 43 37.7 342 2 T23500 hypothetical prote

R:Novotny, W.F.; Girard, T.J.; Miletich, J.P.; Broze Jr., G.J.  
 J. Biol. Chem. 264, 18832-18837, 1989  
 A:Title: Purification and characterization of the lipoprotein-associated coagulation inhibitor  
 A:Reference number: A34315; MUID:90036996; PMID:2553722  
 A:Accession: A34315  
 A:Molecule type: Protein  
 A:Residues: 'XX', 31-33, 'L', 35-50 <NOV>  
 A:Experimental source: plasma  
 R:Pedersen, A.H.; Nordfang, O.; Norris, F.; Wiberg, F.C.; Christensen, P.M.; Moeller, K.  
 J. Biol. Chem. 265, 16786-16793, 1990  
 A:Title: Recombinant human extrinsic pathway inhibitor. Production, isolation, and characterization  
 A:Reference number: A38294; MUID:91009092; PMID:2211593  
 A:Accession: A38294  
 A:Molecule type: protein  
 A:Residues: 29-41 <PED>  
 R:Girard, T.J.; Warren, L.A.; Novotny, W.F.; Likert, K.M.; Brown, S.G.; Miletich, J.P.; Nature 338, 518-520, 1989  
 A:Title: Functional significance of the Kunitz-type inhibitory domains of lipoprotein-associated  
 A:Reference number: S03903; MUID:89181950; PMID:2927510  
 A:Contents: annotation; site-directed mutagenesis  
 C:Comment: The first Kunitz-type domain binds the factor VIIa/tissue factor complex; the  
 C:Genetics:  
 A:Gene: GDB:TFPI  
 A:Cross-references: GDB:127364; OMIM:152310  
 A:Map position: 2q32-2q32  
 A:Introns: 41/1; 107/1; 120/1; 179/1; 210/1; 270/1  
 C:Function:  
 A:Description: regulates clotting by factor Xa-dependent inhibition of the coagulation  
 A:Pathway: blood coagulation  
 C:Superfamily: tissue factor pathway inhibitor; animal Kunitz-type proteinase inhibitor  
 C:Keywords: anticoagulant; blood coagulation; duplication; glycoprotein; heparin binding  
 F:1-28/Domain: signal sequence #status predicted <SIG>  
 F:29-304/Product: tissue factor pathway inhibitor #status experimental <MAT>  
 F:54-104/Domain: animal Kunitz-type proteinase inhibitor homology <BP1>  
 F:125-175/Domain: animal Kunitz-type proteinase inhibitor homology <BP2>  
 F:217-267/Domain: animal Kunitz-type proteinase inhibitor homology <BP3>  
 F:284-289/Region: heparin binding #status predicted  
 F:30/Binding site: phosphate (Ser) (covalent) #status experimental  
 F:54-104, 63-87, 79-100, 125-175, 134-158, 150-171, 217-267, 226-250, 242-263/Disulfide bonds: #  
 F:64/Inhibitory site: Lys (coagulation factor VII/tissue factor complex) #status experimental  
 F:135/Inhibitory site: Arg (coagulation factor X) #status experimental  
 F:145,195,256/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F:227/Inhibitory site: Arg (unidentified proteinase) #status predicted

Query Match 100.0%; Score 114; DB 1; Length 304;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-08;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTKRRKKQKRVKIAYEEIFVKNN 23  
 DB 282 KTKRRKKQKRVKIAYEEIFVKNN 304  
 |||||

RESULT 2  
 JC2264  
 tissue factor pathway inhibitor precursor - rhesus macaque  
 N:Alternate names: extrinsic pathway inhibitor; lipoprotein-associated coagulation inhibitor  
 C:Species: Macaca mulatta (rhesus macaque)  
 C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 21-Jul-2000  
 C:Accession: JC2264  
 R:Kamei, S.; Kamikubo, Y.; Hamuro, T.; Fujimoto, H.; Ishihara, M.; Yonemura, H.; Miyamoto  
 J. Biochem. 115, 708-714, 1994  
 A:Title: Amino acid sequence and inhibitory activity of rhesus monkey tissue factor path  
 A:Reference number: JC2264; MUID:94375417; PMID:8089087  
 A:Accession: JC2264  
 A:Molecule type: mRNA  
 A:Residues: 1-304 <KAM>  
 A:Cross-references: GB:S73337; NID:G685016; PIDN:AAB31955.1; PID:G685017  
 A:Experimental source: liver  
 C:Comment: This protein inhibits the activities of factor Xa and tissue factor-factor VII  
 C:Superfamily: tissue factor pathway inhibitor; animal Kunitz-type proteinase inhibitor  
 C:Keywords: anticoagulant; glycoprotein; serine proteinase inhibitor  
 F:1-28/Domain: signal sequence #status predicted <SIG>

Query Match 100.0%; Score 114; DB 1; Length 304;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-08;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTKRRKKQKRVKIAYEEIFVKNN 23  
 DB 282 KTKRRKKQKRVKIAYEEIFVKNN 304  
 |||||

RESULT 3  
 S53325  
 tissue factor pathway inhibitor - rabbit  
 C:Species: Oryctolagus cuniculus (domestic rabbit)  
 C:Date: 01-Aug-1995 #sequence\_revision 01-Sep-1995 #text\_change 16-Jul-1999  
 C:Accession: S53325  
 R:Girard, T.J.; Gallani, D.; Broze Jr., G.J.  
 Biochem. J. 303, 923-928, 1994  
 A:Title: Complementary DNA sequencing of canine tissue factor pathway inhibitor reveals  
 A:Reference number: S53325; MUID:95071310; PMID:7980463  
 A:Accession: S53325  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-396 <GIR>  
 C:Superfamily: animal Kunitz-type proteinase inhibitor homology  
 C:Keywords: serine proteinase inhibitor  
 F:53-103/Domain: animal Kunitz-type proteinase inhibitor homology <BP1>  
 F:125-175/Domain: animal Kunitz-type proteinase inhibitor homology <BP2>  
 F:309-359/Domain: animal Kunitz-type proteinase inhibitor homology <BP3>

Query Match 79.8%; Score 91; DB 2; Length 396;  
 Best Local Similarity 78.3%; Pred. No. 3.6e-05;  
 Matches 18; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 KTKRRKKQKRVKIAYEEIFVKNN 23  
 DB 374 KTKRRKKQKRVKIAYEEIFVKKL 396  
 |||||

RESULT 4  
 I46937  
 tissue factor pathway inhibitor - rabbit  
 C:Species: Oryctolagus cuniculus (domestic rabbit)  
 C:Date: 04-Sep-1997 #sequence\_revision 04-Sep-1997 #text\_change 13-Aug-1999  
 C:Accession: I46937  
 R:Belaudouj, A.; Kuppaswamy, M.N.; Birktoft, J.J.; Bajaj, S.P.  
 Thromb. Res. 69, 547-553, 1993  
 A:Title: Revised cDNA sequence of rabbit tissue factor pathway inhibitor.  
 A:Reference number: I46937; MUID:93276427; PMID:8503123  
 A:Accession: I46937  
 A:Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-299 <BEL>  
 C:Superfamily: tissue factor pathway inhibitor; animal Kunitz-type proteinase inhibitor  
 F:49-99/Domain: animal Kunitz-type proteinase inhibitor homology <BP1>  
 F:120-170/Domain: animal Kunitz-type proteinase inhibitor homology <BP2>  
 F:212-262/Domain: animal Kunitz-type proteinase inhibitor homology <BP12>

Query Match 67.5%; Score 77; DB 2; Length 299;  
 Best Local Similarity 76.2%; Pred. No. 0.0024;  
 Matches 16; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 KTKRRKKQKRVKIAYEEIFVK 21

Db 277 KTKRKKKQPKVITYVETFK 297  
|||||:|||||:|||||:|||||  
RESULT 5  
S12143  
lipoprotein-associated coagulation inhibitor precursor - rabbit  
N:Alternate names: endothelial cell coagulation inhibitor; endothelial cell tissue factor  
C:Species: Oryctolagus cuniculus (domestic rabbit)  
C>Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 16-Jul-1999  
C:Accession: S12143; A61373  
R:Wesselschmidt, R.L.; Girard, T.J.; Broze Jr., G.J.  
Nucleic Acids Res. 18, 6440, 1990  
A:Title: cDNA sequence of rabbit lipoprotein-associated coagulation inhibitor.  
A:Reference number: S12143; MUID:91057146; PMID:2136251  
A:Accession: S12143  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-300 <WES>  
A:Cross-references: EMBL:X54708; NID:g1612; PIDN:CAA38515.1; PID:g1613  
R:Colburn, P.; Crabb, J.W.; Buonassisi, V.  
J. Cell. Physiol. 148, 320-326, 1991  
A:Title: Enhanced inhibition of tissue factor by the extended form of an endothelial cell  
A:Reference number: A61373; MUID:91349227; PMID:1880157  
A:Accession: A61373  
A:Molecule type: protein  
A:Residues: 25-33,'X',35-46 <COL>  
C:Superfamily: tissue factor pathway inhibitor; animal Kunitz-type proteinase inhibitor  
C:Keywords: anticoagulant; glycoprotein  
F:50-100/Domain: animal Kunitz-type proteinase inhibitor homology <BP1>  
F:121-171/Domain: animal Kunitz-type proteinase inhibitor homology <BP2>  
F:213-263/Domain: animal Kunitz-type proteinase inhibitor homology <BP3>  
Query Match 67.5%; Score 77; DB 2; Length 300;  
Best Local Similarity 76.28; Pred. No. 0.0024; Mismatches 4; Indels 0; Gaps 0;  
Matches 16; Conservative 1;  
QY 1 KTKRKKKQPKVITYVETFK 21  
|||||:|||||:|||||:|||||  
Db 278 KTKRKKKQPKVITYVETFK 298  
RESULT 6  
T01480  
hypothetical protein F1707.4 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C>Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 04-Mar-2000  
C:Accession: T01480  
R:Vyotskaia, V.S.; Schwartz, J.R.; Toriumi, M.; Yu, G.; Kwan, A.; Oji, O.; Liu, S.; Li,  
rtz, D.; Li, Y.; Palm, C.J.; Shinn, P.; Sun, H.; Davis, R.W.; Ecker, J.R.; Federspiel, N  
submitted to the EMBL Data Library, June 1998  
A:Description: Arabidopsis thaliana chromosome 1 BAC F1707 sequence.  
A:Reference number: Z14334  
A:Accession: T01480  
A>Status: translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-272 <VYS>  
A:Cross-references: EMBL:AC003671; NID:g2833627; PID:g3176675; GSPDB:GN00059; ATSP:F1707  
C:Genetics:  
A:Gene: ATSP:F1707.4  
A:Map position: 1  
C:Superfamily: Arabidopsis thaliana hypothetical protein F1707.4  
Query Match 49.1%; Score 56; DB 2; Length 272;  
Best Local Similarity 40.9%; Pred. No. 1.8; Mismatches 8; Indels 0; Gaps 0;  
Matches 9; Conservative 8;  
QY 1 KTKRKKKQPKVITYVETFK 22  
|||||:|||||:|||||:|||||  
Db 217 KTKRKKKQPKVITYVETFK 238  
RESULT 7

T18489  
hypothetical protein C0820w - malaria parasite (Plasmodium falciparum)  
C:Species: Plasmodium falciparum  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jun-2000  
C:Accession: T18489  
R:Lawson, D.; Bowman, S.; Barrell, B.  
submitted to the EMBL Data Library, August 1997  
A:Reference number: Z18935  
A:Accession: T18489  
A>Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-4981 <LAW>  
A:Cross-references: EMBL:Z98551; NID:el331903; PID:el331910; PIDN:CAB11128.1  
C:Genetics:  
A:Map position: 3  
A:Note: C0820w  
Query Match 45.6%; Score 52; DB 2; Length 4981;  
Best Local Similarity 50.0%; Pred. No. 91; Mismatches 4; Indels 0; Gaps 0;  
Matches 11; Conservative 4;  
QY 1 KTKRKKKQPKVITYVETFK 22  
|||||:|||||:|||||:|||||  
Db 4217 KTKRKKKQPKVITYVETFK 4238  
RESULT 8  
C97205  
uncharacterized protein, homolog of Spirochaeta aurantia (gi,152901) [Imported] - CLO  
C:Species: Clostridium acetobutylicum  
C>Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 14-Sep-2001  
C:Accession: C97205  
R:Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; L  
; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.  
J. Bacteriol. 183, 4823-4838, 2001  
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium  
A:Reference number: A96900; MUID:21359325; PMID:21359325  
A:Accession: C97205  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-114 <KUR>  
A:Cross-references: GB:AE001437; PIDN:AAK80430.1; PID:g15025495; GSPDB:GN00168  
A:Experimental source: Clostridium acetobutylicum ATCC824  
C:Genetics:  
A:Gene: CAC2476  
Query Match 43.0%; Score 49; DB 2; Length 114;  
Best Local Similarity 53.3%; Pred. No. 7.6; Mismatches 4; Indels 0; Gaps 0;  
Matches 8; Conservative 4;  
QY 1 KTKRKKKQPKVITYVETFK 15  
|||||:|||||:|||||:|||||  
Db 100 KTKRKKKQPKVITYVETFK 114  
RESULT 9  
G90529  
heat shock protein grpe (activation of DNak) [Imported] - Mycoplasma pulmonis (strain  
C:Species: Mycoplasma pulmonis  
C>Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 03-Aug-2001  
C:Accession: G90529  
R:Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galisson, F.; Moszer,  
Nucleic Acids Res. 29, 2145-2153, 2001  
A:Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma p  
A:Reference number: A99512; MUID:21267165; PMID:11353084  
A:Accession: G90529  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-298 <KUR>  
A:Cross-references: GB:AL445566; PID:g14089556; PIDN:CAC13316.1; GSPDB:GN00153  
C:Genetics:  
A:Experimental source: strain UAB CTIP  
A:Gene: MYPU\_1430







GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 2, 2002, 09:59:55 ; Search time 2.28856 Seconds  
(without alignments)  
416.837 Million cell updates/sec

Title: US-09-741-106-10

Perfect score: 114  
Sequence: 1 KTRKRKKQKVIAEIEFVKNM 23

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | ID           | Description        |
|------------|-------|-------------|--------|--------------|--------------------|
| 1          | 114   | 100.0       | 304    | 1 TFPI_HUMAN | P10646 homo sapien |
| 2          | 113   | 99.1        | 304    | 1 TFPI_MACMU | Q28864 macaca mula |
| 3          | 77    | 67.5        | 300    | 1 TFPI_RABIT | P19761 oryctolagus |
| 4          | 48.5  | 42.5        | 135    | 1 YDQA_SCHPO | O14202 schizosacch |
| 5          | 48    | 42.1        | 727    | 1 YR22_CAEEL | Q09639 caenorhabdi |
| 6          | 48    | 42.1        | 2025   | 1 TTC3_HUMAN | P53804 homo sapien |
| 7          | 47    | 41.2        | 803    | 1 GYRB_BUCAP | P29435 buchnera ap |
| 8          | 46    | 40.4        | 240    | 1 RNC_THEMA  | Q9X016 thermotoga  |
| 9          | 45    | 39.5        | 402    | 1 OXAL_YEAST | P39952 saccharomyc |
| 10         | 45    | 39.5        | 604    | 1 Y426_HUMAN | O43309 homo sapien |
| 11         | 45    | 39.5        | 2210   | 1 RRPO_LYCVA | P14240 lymphocytic |
| 12         | 44.5  | 39.0        | 667    | 1 SYG_YEAST  | P38088 saccharomyc |
| 13         | 44    | 38.6        | 368    | 1 SNCL_HUMAN | Q16533 homo sapien |
| 14         | 43    | 37.7        | 1085   | 1 YAFB_SCHPO | Q09863 schizosacch |
| 15         | 43    | 37.7        | 1142   | 1 KPCL_NEUCR | P87253 neurospora  |
| 16         | 42.5  | 37.3        | 167    | 1 CAT3_HUMAN | O15182 homo sapien |
| 17         | 42.5  | 37.3        | 167    | 1 CAT3_MOUSE | O35648 mus musculu |
| 18         | 42    | 36.8        | 71     | 1 Y647_METJA | Q58063 methanococ  |
| 19         | 42    | 36.8        | 319    | 1 TRXB_BUCAP | P81433 buchnera ap |
| 20         | 42    | 36.8        | 529    | 1 NOPS_HUMAN | Q9V2X3 homo sapien |
| 21         | 42    | 36.8        | 696    | 1 FLHA_BUCAI | P57335 buchnera ap |
| 22         | 42    | 36.8        | 959    | 1 MFD_HELPY  | O26066 helicobacte |
| 23         | 42    | 36.8        | 1001   | 1 MFD_HELPJ  | Q92J57 helicobacte |
| 24         | 42    | 36.8        | 1093   | 1 YK05_CAEEL | Q03563 caenorhabdi |
| 25         | 42    | 36.8        | 1195   | 1 HELS_METJA | Q58524 methanococ  |
| 26         | 42    | 36.8        | 2005   | 1 CIN2_RAT   | P04775 rattus norv |
| 27         | 41.5  | 36.4        | 302    | 1 TFPI_RAT   | Q02445 rattus norv |
| 28         | 41.5  | 36.4        | 1311   | 1 ATRX_DROME | Q99qn5 drosophila  |
| 29         | 41    | 36.0        | 86     | 1 RL23_PPRAB | Q9vit7 pyrococcus  |
| 30         | 41    | 36.0        | 118    | 1 VAGL_BOVIN | P79251 bos taurus  |
| 31         | 41    | 36.0        | 118    | 1 VAGL_HUMAN | O75348 homo sapien |
| 32         | 41    | 36.0        | 118    | 1 VAGL_MOUSE | Q9cr51 mus musculu |
| 33         | 41    | 36.0        | 196    | 1 RL19_HUMAN | P14118 homo sapien |

## ALIGNMENTS

## RESULT 1

| ID | TFPI_HUMAN                                                             | STANDARD; | PRT; | 304 AA. |
|----|------------------------------------------------------------------------|-----------|------|---------|
| AC | P10646: O95103;                                                        |           |      |         |
| DT | 01-JUL-1989 (Rel. 11, Created)                                         |           |      |         |
| DT | 01-JUL-1989 (Rel. 11, Last sequence update)                            |           |      |         |
| DT | 15-JUN-2002 (Rel. 41, Last annotation update)                          |           |      |         |
| DE | Tissue factor pathway inhibitor precursor (TFPI) (Lipoprotein-         |           |      |         |
| DE | associated coagulation inhibitor) (LACI) (Extrinsic pathway inhibitor) |           |      |         |
| DE | (EPI).                                                                 |           |      |         |
| GN | TFPI OR TFPII OR LACI.                                                 |           |      |         |
| OS | Homo sapiens (Human).                                                  |           |      |         |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;      |           |      |         |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.             |           |      |         |
| OX | NCBI_TaxID=9606;                                                       |           |      |         |
| RN | [1]                                                                    |           |      |         |
| RP | SEQUENCE FROM N.A. (ISOFORM ALPHA).                                    |           |      |         |
| RX | MEDLINE=88198127; PubMed=2452157;                                      |           |      |         |
| RA | Wun T.-C., Kretzmer K.K., Girard T.J., Miletich J.P., Broze G.J. Jr.;  |           |      |         |
| RT | "Cloning and characterization of a cDNA coding for the lipoprotein-    |           |      |         |
| RT | associated coagulation inhibitor shows that it consists of three       |           |      |         |
| RT | tandem Kunitz-type inhibitory domains.";                               |           |      |         |
| RL | J. Biol. Chem. 263:6001-6004(1988).                                    |           |      |         |
| RN | [2]                                                                    |           |      |         |
| RP | SEQUENCE FROM N.A. (ISOFORM ALPHA).                                    |           |      |         |
| RX | MEDLINE=91129227; PubMed=1993173;                                      |           |      |         |
| RA | van der Logt C.P.E., Reitsma P.H., Bertina R.M.;                       |           |      |         |
| RT | "Intron-exon organization of the human gene coding for the             |           |      |         |
| RT | lipoprotein-associated coagulation inhibitor: the factor Xa dependent  |           |      |         |
| RT | inhibitor of the extrinsic pathway of coagulation.";                   |           |      |         |
| RL | Biochemistry 30:1571-1577(1991).                                       |           |      |         |
| RN | [3]                                                                    |           |      |         |
| RP | SEQUENCE FROM N.A. (ISOFORM ALPHA).                                    |           |      |         |
| RX | MEDLINE=91161593; PubMed=2002045;                                      |           |      |         |
| RA | Girard T.J., Eddy R., Wesselschmidt R.L., Macphail L.A.,               |           |      |         |
| RA | Likert K.M., Byers M.G., Shows T.B., Broze G.J. Jr.;                   |           |      |         |
| RT | "Structure of the human lipoprotein-associated coagulation inhibitor   |           |      |         |
| RT | gene. Intro/exon gene organization and localization of the gene to     |           |      |         |
| RT | chromosome 2.";                                                        |           |      |         |
| RL | J. Biol. Chem. 266:5036-5041(1991).                                    |           |      |         |
| RN | [4]                                                                    |           |      |         |
| RP | SEQUENCE FROM N.A. (ISOFORM ALPHA).                                    |           |      |         |
| RX | MEDLINE=89388722; PubMed=2781520;                                      |           |      |         |
| RA | Girard T.J., Warren L.A., Novotny W.F., Bejcek B.E., Miletich J.P.,    |           |      |         |
| RA | Broze G.J. Jr.;                                                        |           |      |         |
| RT | "Identification of the 1.4 kb and 4.0 kb messages for the lipoprotein  |           |      |         |
| RT | associated coagulation inhibitor and expression of the encoded         |           |      |         |
| RT | protein.";                                                             |           |      |         |
| RL | Thromb. Res. 55:37-50(1989).                                           |           |      |         |
| RN | [5]                                                                    |           |      |         |
| RP | SEQUENCE FROM N.A. (ISOFORM BETA).                                     |           |      |         |
| RA | Chang J.-Y., Monroe D.M., Roberts H.R.;                                |           |      |         |
| RL | Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.                |           |      |         |
| RN | [6]                                                                    |           |      |         |
| RP | SEQUENCE FROM N.A. (ISOFORM BETA).                                     |           |      |         |
| RC | TISSUE=Pancreas;                                                       |           |      |         |

Q9chu2 lactococcus  
P10962 saccharomyc  
Q9y6k9 h ni-kappab  
O14640 homo sapien  
P54792 homo sapien  
Q9gx72 rattus norv  
Q06278 homo sapien  
P23116 mus musculu  
Q99250 homo sapien  
Q8rie4 fusobacteri  
P20547 vaccinia vi  
O26112 methanobact

34 41 36.0 294 1 MIRA\_LACLA  
35 41 36.0 306 1 MK16\_YEAST  
36 41 36.0 419 1 NEMO\_HUMAN  
37 41 36.0 670 1 DVLI\_HUMAN  
38 41 36.0 670 1 DVLI\_HUMAN  
39 41 36.0 846 1 SBP2\_RAT  
40 41 36.0 1338 1 ADO\_HUMAN  
41 41 36.0 1344 1 IF3A\_MOUSE  
42 41 36.0 2005 1 CIN2\_HUMAN  
43 40 35.1 72 1 RS18\_FUSNN  
44 40 35.1 91 1 YV8G\_VACCC  
45 40 35.1 100 1 RL23\_METTH

Strausberg R.;  
Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
[7]  
SEQUENCE OF 29-50.  
MEDLINE=90036996; PubMed=2553722;  
Novotny W.F., Girard T.J., Miletich J.P., Broze G.J. Jr.;  
"Purification and characterization of the lipoprotein-associated  
coagulation inhibitor from human plasma."  
J. Biol. Chem. 264:18832-18837(1989).  
[8]  
INHIBITORY SITES.  
MEDLINE=89181950; PubMed=2927510;  
Girard T.J., Warren L.A., Novotny W.F., Likert K.M., Brown S.G.,  
Miletich J.P., Broze G.J. Jr.;  
"Functional significance of the Kunitz-type inhibitory domains of  
lipoprotein-associated coagulation inhibitor."  
Nature 338:518-520(1989).  
[9]  
CARBOHYDRATE-LINKAGE SITES.  
MEDLINE=96224851; PubMed=86399592;  
Nakahara Y., Miyata T., Hamuro T., Funatsu A., Miyagi M.,  
Tsunawasa S., Kato H.;  
"Amino acid sequence and carbohydrate structure of a recombinant  
human tissue factor pathway inhibitor expressed in Chinese hamster  
ovary cells: one N- and two O-linked carbohydrate chains are located  
between Kunitz domains 2 and 3 and one N-linked carbohydrate chain is  
in Kunitz domain 2."  
Biochemistry 35:6450-6459(1996).  
[10]  
REVIEW.  
MEDLINE=91104709; PubMed=2271516;  
Broze G.J. Jr., Girard T.J., Novotny W.F.;  
"Regulation of coagulation by a multivalent Kunitz-type inhibitor."  
Biochemistry 29:7539-7546(1990).  
[11]  
STRUCTURE BY NMR OF 121-182.  
MEDLINE=97342711; PubMed=9199408;  
Burgering M.J., Orbons L.P., van der Doelen A., Mulders J.,  
Theunissen H.J., Grootenhuis P.D., Bode W., Huber R., Stubbs M.T.;  
"The second Kunitz domain of human tissue factor pathway inhibitor:  
cloning, structure determination and interaction with factor Xa."  
J. Mol. Biol. 269:395-407(1997).  
-1- FUNCTION: INHIBITS FACTOR X (X(A)) DIRECTLY AND, IN A XA-DEPENDENT  
WAY, INHIBITS VII(A)/TISSUE FACTOR ACTIVITY, PRESUMABLY BY FORMING  
A QUATERNARY X(A)/LACI/VII(A)/TF COMPLEX. IT POSSESSES AN  
ANTITHROMBOTIC ACTION AND ALSO THE ABILITY TO ASSOCIATE WITH  
LIPOPROTEIN IN PLASMA.  
-1- SUBCELLULAR LOCATION: Secreted.  
-1- ALTERNATIVE PRODUCTS: 2 isoforms; alpha/TFPIalpha (shown here) and  
beta/TFPIbeta; are produced by alternative splicing.  
-1- TISSUE SPECIFICITY: MOSTLY IN ENDOTHELIAL CELLS.  
-1- DOMAIN: THIS INHIBITOR CONTAINS THREE INHIBITORY DOMAINS.  
-1- PTM: O-GLYCOSYLATED.  
-1- SIMILARITY: CONTAINS 3 BPTI/KUNITZ INHIBITOR DOMAINS.  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
EMBL; J03225; AAA52022.1; -  
EMBL; M58650; AAA59480.1; -  
EMBL; M58644; AAA59480.1; JOINED.  
EMBL; M58645; AAA59480.1; JOINED.  
EMBL; M58646; AAA59480.1; JOINED.  
EMBL; M58647; AAA59480.1; JOINED.  
EMBL; M58648; AAA59480.1; JOINED.  
EMBL; M58649; AAA59480.1; JOINED.  
EMBL; M59499; AAA59526.1; -  
EMBL; M59493; AAA59526.1; JOINED.





DR PROSITE; PS02079; BPTI\_KUNITZ\_2; 3.  
 KW Serine protease inhibitor; Glycoprotein; Repeat; Blood coagulation;  
 FT SIGNAL. 1 24  
 FT CHAIN 25 300  
 FT DOMAIN 50 100  
 TISSUE FACTOR PATHWAY INHIBITOR.  
 BPTI/KUNITZ INHIBITOR 1  
 (VII(A)/TISSUE FACTOR BINDING SITE).  
 BPTI/KUNITZ INHIBITOR 2  
 (FACTOR X(A) BINDING SITE).  
 BPTI/KUNITZ INHIBITOR 3  
 (BY SIMILARITY).  
 FT DOMAIN 213 263  
 FT DISULFID 50 100  
 FT DISULFID 59 83  
 FT DISULFID 75 96  
 FT ACT\_SITE 60 61  
 FT DISULFID 121 171  
 FT DISULFID 130 154  
 FT DISULFID 146 167  
 FT ACT\_SITE 131 132  
 FT DISULFID 213 263  
 FT DISULFID 222 246  
 FT DISULFID 238 259  
 FT ACT\_SITE 223 224  
 FT CARBOHYD 141 141  
 FT CARBOHYD 191 191  
 FT CARBOHYD 252 252  
 FT CONFLICT 31 31  
 FT CONFLICT 269 272  
 FT CONFLICT PKSI -> RNLS (IN REF. 3).  
 FT SEQUENCE 300 AA; 34435 MW; A08DE36537708CA6 CRC64;  
 Query Match 67.5%; Score 77; DB 1; Length 300;  
 Best Local Similarity 76.2%; Pred. No. 0.00033;  
 Matches 16; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 KTKRRKKRKQVKTAYEEIFVK 21  
 DB 278 KTKRRKKRKQVKTAYEEIFVK 298  
 RESULT 4  
 ID YDQA\_SCHPO STANDARD; PRT; 135 AA.  
 AC 014202;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Hypothetical protein C5D6.10c in chromosome I.  
 GN SPAC5D6.10c.  
 OS Schizosaccharomyces pombe (Fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomycetes.  
 OX NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972;  
 RX MEDLINE=21848401; PubMed=11859360;  
 RA Wood V., Williams R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
 RA Sgourou J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,  
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,  
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,  
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
 RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,  
 RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
 RA Gabel C., Fuchs M., Fritz C., Holzer M., Moestl D., Hilbert H.,  
 RA Borzym K., Langer I., Beck A., Leinhardt R., Reinhardt T., Pohl T.M.,

RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,  
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,  
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
 RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
 RA Shpakowski G.V., Ussery D., Barrell B.G., Nurse P.;  
 RT "The genome sequence of Schizosaccharomyces pombe";  
 RL Nature 415:871-880(2002).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC DR EMBL; Z98056; CAB10858.1; -.  
 KW Hypothetical protein.  
 FT DOMAIN 85 111 LYS-RICH.  
 SQ SEQUENCE 135 AA; 15864 MW; 2F39A1F4A909476D CRC64;  
 Query Match 42.5%; Score 48.5; DB 1; Length 135;  
 Best Local Similarity 54.5%; Pred. No. 2;  
 Matches 12; Conservative 3; Mismatches 4; Indels 3; Gaps 1;  
 QY 1 KTKRRKKRKQVKTAYEEIFVK 22  
 DB 89 KTKRRKKRKQVKTAYEEIFVK 107  
 RESULT 5  
 YR22\_CAEEL STANDARD; PRT; 727 AA.  
 AC Q09639;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Putative G protein-coupled receptor kinase W02B3.2 (EC 2.7.1.-).  
 GN W02B3.2.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Connell M.;  
 RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.  
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 CC GPRK SUBFAMILY.  
 CC -!- SIMILARITY: CONTAINS 1 RGS DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC DR EMBL; U22833; AAA64322.1; -.  
 KW HSP; P25098; 1BAK  
 DR WormPep; W02B3.2; CE02074.  
 DR InterPro; IPR000719; Euk\_pkinase.  
 DR InterPro; IPR000239; GPCR\_kinase.  
 DR InterPro; IPR001849; PH.  
 DR InterPro; IPR000961; pkinase\_C.  
 DR InterPro; IPR000342; Regl\_Gproteins.  
 DR InterPro; IPR002290; Ser\_thr\_pkinase.  
 DR Pfam; PF00069; pkinase; 1.

DR pfam: PF00615; RGS; 1.  
 DR PRINTS: PR00717; GPCR\_KINASE.  
 DR ProDom; PD000001; Euk\_pkinase; 1.  
 DR SMART; SM00233; PH; 1.  
 DR SMART; SM00315; RGS; 1.  
 DR SMART; SM00133; S-TK\_X; 1.  
 DR SMART; SM00220; S-TKC; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS00111; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 DR PROSITE; PS00132; RGS; 1.  
 KW Hypothetical protein; Transferase; Serine/threonine-protein kinase;  
 KW ATP-binding.  
 FT DOMAIN 1 190 N-TERMINAL.  
 FT DOMAIN 191 455 PROTEIN KINASE.  
 FT DOMAIN 456 700 C-TERMINAL.  
 FT DOMAIN 54 175 RGS.  
 FT NP\_BIND 197 205 ATP (BY SIMILARITY).  
 FT BINDING 220 220 ATP (BY SIMILARITY).  
 FT ACT\_SITE 318 318 BY SIMILARITY.  
 SQ SEQUENCE 727 AA; 83361 MW; AA0530D8A9DA57D7 CRC64;  
 Query Match 42.1%; Score 48; DB 1; Length 727;  
 Best Local Similarity 57.1%; Pred. No. 14;  
 Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;  
 QY 4 RRRKQKRVKAYEE 17  
 Db 541 KRRSQKIKVAVEE 554  
 :|||:|:|:|:|:|  
 RESULT 6  
 ID TTC3\_HUMAN STANDARD; PRT; 2025 AA.  
 AC P53804; P78476; P78477;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Tetratricopeptide repeat protein 3 (TPR repeat protein D).  
 GN TTC3 OR TPRD.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=96281123; PubMed=8724848;  
 RA Ohira M., Ootsuyama A., Suzuki E., Ichikawa H., Seki N.,  
 RA Nagase T., Nomura N., Ohki M.;  
 RT "Identification of a novel human gene containing the  
 RT tetratricopeptide repeat domain from the Down syndrome region of  
 RT chromosome 21.";  
 RL DNA Res. 3:9-16(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Fetal brain;  
 RX MEDLINE=97103476; PubMed=8947847;  
 RA Tsukahara F., Hattori M., Muraki T., Sakaki Y.;  
 RT "Identification and cloning of a novel cDNA belonging to  
 RT tetratricopeptide repeat gene family from Down syndrome-critical  
 RT region 21q22.2.";  
 RL J. Biochem. 120:820-827(1996).  
 CC -!- ALTERNATIVE PRODUCTS: 3 ISOFORMS; TRPDI (SHOWN HERE), TPRDII  
 CC AND TPRDIII; SEEMS TO BE PRODUCED BY ALTERNATIVE SPLICING.  
 CC -!- TISSUE SPECIFICITY: FOUND IN ALL TISSUES EXAMINED.  
 CC -!- SIMILARITY: CONTAINS 4 TPR REPEATS.  
 CC -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 DR EMBL; D83077; BAA11769.1; -;  
 DR EMBL; D84294; BAA12301.1; -;  
 DR EMBL; D84295; BAA12302.1; -;  
 DR EMBL; D84296; BAA12303.1; -;  
 DR Genew; HGNC:12393; TTC3.  
 DR MIM; 602259; -;  
 DR InterPro; IPR001440; TPR.  
 DR InterPro; IPR001841; Znf\_ring.  
 DR Pfam; PF00097; zf-C3HC4; 1.  
 DR Pfam; PF00515; TPR; 3.  
 DR SMART; SM00184; RING; 1.  
 DR SMART; SM00028; TPR; 3.  
 DR PROSITE; PS00518; ZF\_RING\_1; FALSE\_NEG.  
 DR PROSITE; PS00089; ZF\_RING\_2; 1.  
 KW Repeat; TPR repeat; Zinc-finger; Alternative splicing.  
 FT REPEAT 231 264 TPR 1.  
 FT REPEAT 266 298 TPR 2.  
 FT REPEAT 536 572 TPR 3.  
 FT REPEAT 576 609 TPR 4.  
 FT ZN\_FING 1957 1997 RING-TYPE.  
 FT DOMAIN 453 456 POLY-SER.  
 FT DOMAIN 1020 1029 POLY-LYS.  
 FT DOMAIN 1899 1902 POLY-LYS.  
 FT DOMAIN 1018 1029 ARG/LYS-RICH (BASIC).  
 FT DOMAIN 1172 1185 ARG/LYS-RICH (BASIC).  
 FT DOMAIN 1563 1579 ARG/LYS-RICH (BASIC).  
 FT VARSPLIC 1 233 MISSING (IN ISOFORM TPRDII).  
 FT VARSPLIC 1 310 MISSING (IN ISOFORM TPRDIII).  
 SQ SEQUENCE 2025 AA; 229889 MW; 1B4BCAA3684B6253 CRC64;  
 Query Match 42.1%; Score 48; DB 1; Length 2025;  
 Best Local Similarity 50.0%; Pred. No. 41;  
 Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;  
 QY 1 KTKRKRKKQKRVKAYEEI 18  
 Db 1177 RLKKRKRKKKIKTKVVEI 1194  
 :|:|:|:|:|:|:|:|  
 RESULT 7  
 ID GYRB\_BUCAP STANDARD; PRT; 803 AA.  
 AC P29435;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE DNA gyrase subunit B (EC 5.99.1.3).  
 GN GYRB.  
 OS Buchnera aphidicola (subsp. Schizaphis graminum).  
 OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.  
 OX NCBI\_TaxID=98794;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98184963; PubMed=9516544;  
 RA Clark M.A., Baumann L., Baumann P.;  
 RT "Sequence analysis of a 34.7-kb DNA segment from the genome of  
 RT Buchnera aphidicola (endosymbiont of aphids) containing groEL, dnaA,  
 RT the atp operon, gidA, and rbo.";  
 RL Curr. Microbiol. 36:158-163(1998).  
 RN [2]  
 RP SEQUENCE OF 1-91 FROM N.A.  
 RX MEDLINE=92241666; PubMed=1572539;  
 RA Lai C.-Y., Baumann P.;  
 RT "Genetic analysis of an aphid endosymbiont DNA fragment homologous to  
 RT the rnpA-rnpH-dnaA-dnaN-gyrB region of eubacteria.";  
 RL Gene 113:175-181(1992).  
 CC -!- FUNCTION: DNA GYRASE NEGATIVELY SUPERCOILS CLOSED CIRCULAR DOUBLE-  
 CC STRANDED DNA IN AN ATP-DEPENDENT MANNER AND ALSO CATALYZES THE  
 CC INTERCONVERSION OF OTHER TOPOLOGICAL ISOMERS OF DOUBLE-STRANDED

```

CC DNA RINGS, INCLUDING CATENANES AND KNOTTED RINGS.
CC -!- CATALYTIC ACTIVITY: ATP-dependent breakage, passage and rejoining
CC of double-stranded DNA.
CC -!- SUBUNIT: MADE UP OF TWO CHAINS. THE A CHAIN IS RESPONSIBLE FOR DNA
CC BREAKAGE AND REJOINING; THE B CHAIN CATALYZES ATP HYDROLYSIS. THE
CC ENZYME FORMS AN A2B2 TETRAMER.
CC -!- SIMILARITY: BELONGS TO THE TYPE II TOPOISOMERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF008210; AAC38108.1; -
CC EMBL: M80817; AAA73151.1; -
CC FIC: JCI160; JCI160.
CC HSP: P06982; IAJ6.
CC InterPro: IPR003594; ATPbind_ATPase.
CC InterPro: IPR002288; DNA_gyraseB_C.
CC InterPro: IPR001241; DNA_topoisomII.
CC InterPro: IPR002336; DNAPrim_toprim.
CC Pfam: PF00204; DNA_gyraseB; 1.
CC Pfam: PF00986; DNA_gyraseB_C; 1.
CC Pfam: PF01751; Toprim; 1.
CC Pfam: PF02518; HATPase_C; 1.
CC PRINTS: PR00418; TP12FAMILY.
CC ProDom: PD000616; DNA_topoisomII; 1.
CC ProDom: PD149633; DNA_gyraseB_C; 1.
CC SMART: SM00387; HATPase_C; 1.
CC SMART: SM00433; TOP2c; 1.
CC TIGRFAMs: TIGR01059; gyrB; 1.
CC PROSITE: PS00177; TOPOISOMERASE-II; 1.
KW Topoisomerase; Isomerase; ATP-binding.
SQ SEQUENCE 803 AA; 92678 MW; C88758E8AB5FE6D2 CRC64;

Query Match 41.2%; Score 47; DB 1; Length 803;
Best Local Similarity 40.9%; Pred. No. 22;
Matches 9; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 1 KTKRRKKQKVKIAYEEFVKV 22
DB 547 KNDEEMNKYQIKALKKEIVKN 568

RESULT 8
RNC_THEME
ID RNC_THEME STANDARD; PRT; 240 AA.
AC Q9X016;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ribonuclease III (EC 3.1.26.3) (Rnase III).
GN RNC OR TM1102.
OS Thermotoga maritima.
OC Bacteria; Thermotogae; Thermotogae (class); Thermotogales;
OC Thermotogaceae; Thermotoga.
OX NCBI_TaxID=2336;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MSB8 / DSM 3109;
RX MEDLINE=99287316; PubMed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RA "Evidence for lateral gene transfer between Archaea and Bacteria from
RT genome sequence of Thermotoga maritima.";
RL Nature 399:323-329(1999).

```

```

CC -!- FUNCTION: DIGESTS DOUBLE-STRANDED RNA. INVOLVED IN THE PROCESSING
CC OF RIBOSOMAL RNA PRECURSORS AND OF SOME MRNAs (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphomonoester.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: CONTAINS 1 DRBM (DOUBLE-STRANDED RNA-BINDING) DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 RNASE III DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE001769; AAD36178.1; -
CC TIGR: TM1102; -
CC InterPro: IPR001159; DS_RBD.
CC InterPro: IPR000999; RNase_3.
CC Pfam: PF00035; dsrm; 1.
CC Pfam: PF00636; Ribonuclease_3; 1.
CC SMART: SM00358; DSRM; 1.
CC SMART: SM00535; RIBOC; 1.
CC PROSITE: PS00137; DS_RBD; 1.
CC PROSITE: PS00517; RNase_3_1; 1.
CC PROSITE: PS0142; RNase_3_2; 1.
KW Hydrolase; Nuclease; Endonuclease; RNA-binding; Complete proteome.
FT DOMAIN 9 141 RNase III.
FT DOMAIN 219 235 DRBM.
SQ SEQUENCE 240 AA; 27530 MW; 94330E8898D48A0D CRC64;

Query Match 40.4%; Score 46; DB 1; Length 240;
Best Local Similarity 38.9%; Pred. No. 8;
Matches 7; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

QY 1 KTKRRKKQKVKIAYEEI 18
DB 218 RTKKEAEKEAARIAYEKL 235

RESULT 9
OXAL_YEAST
ID OXAL_YEAST STANDARD; PRT; 402 AA.
AC P39952;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Cytochrome oxidase biogenesis protein OXAL, mitochondrial precursor.
GN OXAL OR PET1402 OR YER154W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FL1100;
RX MEDLINE=94254098; PubMed=8196054;
RA Bonnefoy N., Chalvet F., Hamel P., Slonimski P.P., Dujardin G.;
RT "OXAL, a Saccharomyces cerevisiae nuclear gene whose sequence is
RT conserved from prokaryotes to eukaryotes controls cytochrome oxidase
RT biogenesis.";
RL J. Mol. Biol. 239:201-212(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RA Dietrich F.S., Mulligan J.T., Hennessey K.M., Allen E., Araujo R.,
RA Aviles E., Berno A., Brennan T., Carpenter J., Chen E., Cherry J.M.,
RA Chung E., Duncan M., Guzman E., Hartzell G., Hunnicke-Smith S.,
RA Hyman R., Kayser A., Komp C., Lashkari D., Lew H., Lin D.,
RA Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P., Oh C.,
RA Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T., Smith V.,
RA Taylor P., Wei Y., Yelton M., Botstein D., Davis R.W.;

```

Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.  
[3]  
SEQUENCE FROM N.A.  
MEDLINE=95115677; PubMed=7816036;  
Bauer M., Behrens M., Esser K., Michaelis G., Pratz E.;  
"PET1402, a nuclear gene required for proteolytic processing of  
cytochrome oxidase subunit 2 in yeast.";  
Mol. Gen. Genet. 245:272-278(1994).  
[4]  
TOPOLOGY.  
MEDLINE=97315163; PubMed=9171337;  
Herrmann J.M., Neupert W., Stuart R.A.;  
"Insertion into the mitochondrial inner membrane of a polytopic  
protein, the nuclear-encoded Oxalp.";  
EMBO J. 16:2217-2226(1997).  
CC -!- FUNCTION: REQUIRED FOR POST-TRANSLATIONAL STEP OF CYTOCHROME  
OXIDASE BIOGENESIS. SEEMS TO BE INVOLVED IN THE PROTEOLYTIC  
PROCESSING OF CYTOCHROME OXIDASE SUBUNIT 2.  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial.  
CC -!- SIMILARITY: BELONGS TO THE OXAL / 60 KDA IMP FAMILY.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See http://www.isb-sib.ch/announce/  
or send an email to license@isb-sib.ch).  
-----  
DR EMBL; X77558; CAA54675.1; -;  
DR EMBL; U18917; AAB64681.1; -;  
DR EMBL; X74456; CAA52465.1; -;  
DR PIR; S47329; S47329.  
DR SGD; S0000956; OXAL.  
DR InterPro; IPR001708; 60kDa\_innermem.  
DR Pfam; PF02096; 60KD\_IMP; 1.  
KW Transmembrane, Mitochondrion; Transit peptide.  
FT CHAIN 1 42 MITOCHONDRION (POTENTIAL).  
FT CHAIN 43 402 CYTOCHROME OXIDASE BIOGENESIS PROTEIN  
Oxal.  
FT DOMAIN 43 130 MITOCHONDRIAL INTERMEMBRANE (PROBABLE).  
FT TRANSMEM 131 148 PROBABLE.  
FT DOMAIN 149 200 MITOCHONDRIAL MATRIX (PROBABLE).  
FT TRANSMEM 201 218 PROBABLE.  
FT DOMAIN 219 246 MITOCHONDRIAL INTERMEMBRANE (PROBABLE).  
FT TRANSMEM 247 263 PROBABLE.  
FT DOMAIN 264 275 MITOCHONDRIAL MATRIX (PROBABLE).  
FT TRANSMEM 276 291 PROBABLE.  
FT DOMAIN 292 292 MITOCHONDRIAL INTERMEMBRANE (PROBABLE).  
FT TRANSMEM 293 310 PROBABLE.  
FT DOMAIN 311 402 MITOCHONDRIAL MATRIX (PROBABLE).  
FT CONFLICT 108 108 W -> R (IN REF. 1).  
SQ SEQUENCE 402 AA; 44815 MW; 108890B8F2EE67ED CRC64;  
  
Query Match 39.5%; Score 45; DB 1; Length 402;  
Best Local Similarity 35.0%; Pred. No. 21;  
Matches 7; Conservative 6; Mismatches 7; Indels 0; Gaps 0;  
  
QY 3 KRKRKQKRVKIAEEIFVKV 22  
| | | : | | : | |  
Db 381 KEKQNSKIKIVHKSNNFN 400  
  
RESULT 10  
Y426\_HUMAN  
ID Y426\_HUMAN STANDARD; PRT; 604 AA.  
AC O43309; O43724;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Hypothetical zinc finger protein KIAA0426.  
GN KIAA0426.  
OS Homo sapiens (Human).  
  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RN SEQUENCE FROM N.A.  
RP TISSUE=Brain;  
RX MEDLINE=98116655; PubMed=9455477;  
RA Tshikawa K.-I., Nagase T., Nakajima D., Seki N., Ohira M.,  
RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;  
"Prediction of the coding sequences of unidentified human genes. VIII.  
78 new cDNA clones from brain which code for large proteins in  
vitro.";  
RNA Res. 4:307-313(1997).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Tubby B.;  
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: MAY BE INVOLVED IN TRANSCRIPTIONAL REGULATION.  
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).  
CC -!- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-  
FINGER PROTEINS.  
CC -!- SIMILARITY: CONTAINS 1 SCAN BOX.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See http://www.isb-sib.ch/announce/  
or send an email to license@isb-sib.ch).  
-----  
DR EMBL; AB007886; BAA24856.1; -;  
DR EMBL; Z98745; CAB11428.2; -;  
DR HSSP; P07248; 2ADR.  
DR InterPro; IPR003309; Treg\_SCAN.  
DR InterPro; IPR000822; Znf\_C2H2.  
DR Pfam; PF00096; zf-C2H2; 10.  
DR Pfam; PF02023; SCAN; 1.  
DR PRINTS; PR00048; ZINC\_FINGER.  
DR CYTOCHROME OXIDASE BIOGENESIS PROTEIN.  
DR PRODOM; PD000003; Znf\_C2H2; 8.  
DR SMART; SM00431; LER; 1.  
DR SMART; SM00355; Znf\_C2H2; 10.  
DR PROSITE; PS00804; SCAN\_BOX; 1.  
DR PROSITE; PS00028; ZINC\_FINGER\_C2H2\_1; 9.  
DR PROSITE; PS0157; ZINC\_FINGER\_C2H2\_2; 11.  
KW Hypothetical protein; Transcription regulation; DNA-binding;  
KW Zinc-finger; Metal-binding; Nuclear protein; Repeat.  
FT DOMAIN 46 128 SCAN\_BOX.  
FT DOMAIN 274 512 ZINC\_FINGERS.  
FT ZN\_FING 274 296 C2H2-TYPE.  
FT ZN\_FING 302 324 C2H2-TYPE.  
FT ZN\_FING 330 352 C2H2-TYPE.  
FT ZN\_FING 358 380 C2H2-TYPE.  
FT ZN\_FING 386 408 C2H2-TYPE.  
FT ZN\_FING 414 436 C2H2-TYPE.  
FT ZN\_FING 442 463 C2H2-TYPE.  
FT ZN\_FING 469 491 C2H2-TYPE.  
FT ZN\_FING 497 519 C2H2-TYPE.  
FT ZN\_FING 525 547 C2H2-TYPE.  
SQ SEQUENCE 604 AA; 70222 MW; 6A7F8ECB515A7A73 CRC64;  
  
Query Match 39.5%; Score 45; DB 1; Length 604;  
Best Local Similarity 50.0%; Pred. No. 32;  
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;  
  
QY 5 KRKRKQKRVKIAEEIFVKV 22  
| | | : | | : | |  
Db 560 KAFRQKTSYKEILLKN 577  
  
RESULT 11  
RRPO\_LYCVA  
ID RRPO\_LYCVA STANDARD; PRT; 2210 AA.

```
AC P14240;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
LN RNA polymerase (EC 2.7.7.48).
GN L.
OS Lymphocytic choriomeningitis virus (strain Armstrong).
OC Viruses: ssRNA negative-strand viruses; Arenaviridae; Arenavirus.
OX NCBI_TaxID=11624;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89204909; PubMed=2705303;
RA Salvato M.S., Shimomaye E.M., Oldstone M.B.A.;
RT "The primary structure of the lymphocytic choriomeningitis virus L
RT gene encodes a putative RNA polymerase.";
RL Virology 169:377-384(1989).
RN [2]
RP SEQUENCE OF 161-387; 424-619 AND 1646-1906 FROM N.A.
RX MEDLINE=88072084; PubMed=3318094;
RA Singh M.K., Fuller-Pace F.V., Buchmeier M.J., Southern P.J.;
RT "Analysis of the genomic L RNA segment from lymphocytic
RT choriomeningitis virus.";
RL Virology 161:448-456(1987).
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate +
CC [RNA]N.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; J04331; AAA66591.1; -;
DR EMBL; M18381; AAA46258.1; ALT_SEQ.
DR EMBL; M18382; AAA46259.1; -;
DR EMBL; M18383; AAA46260.1; ALT_SEQ.
DR PIR; A30181; RXPLC.
KW RNA-directed RNA polymerase; Transferase.
FT CONFLICT 164 164 L -> Y (IN REF. 2).
FT CONFLICT 354 354 Q -> R (IN REF. 2).
FT CONFLICT 361 361 K -> E (IN REF. 2).
FT CONFLICT 382 382 H -> D (IN REF. 2).
FT CONFLICT 552 552 C -> S (IN REF. 2).
FT CONFLICT 1727 1727 R -> L (IN REF. 2).
SQ SEQUENCE 2210 AA; 254529 MW; 470C8E623176AFD3 CRC64;

Query Match 39.5%; Score 45; DB 1; Length 2210;
Best Local Similarity 50.0%; Pred. No. 1.2e+02;
Matches 9; Conservative 5; Mismatches 2; Indels 2; Gaps 1;

QY 4 RRRKKQKV--KIAVEEIF 19
Db 393 RNRKSKVKFRKVSFEELF 410
| ||| :| :|:|:|:|

RESULT 12
SYG_YEAST STANDARD; PRT; 667 AA.
AC P38088;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE GLYCyl-tRNA synthetase (EC 6.1.1.14) (Glycine--tRNA ligase) (GLYRS).
GN GRS1 OR YBR121C OR YBR0917.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c;
```

```
RX MEDLINE=95208357; PubMed=7900426;
RA Mannhaupt G., Stucka R., Ehmler S., Vetter I., Feldmann H.;
RT "Analysis of a 70 kb region on the right arm of yeast chromosome II.";
RL Yeast 10:1363-1381(1994).
CC -!- CATALYTIC ACTIVITY: ATP + glycine + tRNA(Gly) -> AMP + diphosphate
CC + glycyL-tRNA(Gly).
CC -!- SUBUNIT: HOMODIMER.
CC -!- SIMILARITY: BELONGS TO CLASS-II AMINOACYL-TRNA SYNTHETASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; X78993; CAA55623.1; -;
DR EMBL; Z35990; CAA85078.1; -;
DR PIR; S44700; S44700.
DR HSSP; P56206; 1B76.
DR SGD; S0000325; GRS1.
DR InterPro; IPR002106; AATRNA_LigaseII.
DR InterPro; IPR004154; HGTP_anticonodon.
DR InterPro; IPR002314; tRNA-synt_2b.
DR InterPro; IPR002315; tRNA-synt_gly.
DR Pfam; PF00367; tRNA-synt_2b; 1.
DR Pfam; PF03129; HGTP_anticonodon; 1.
DR PRINTS; PR01043; TRNASYNTHGLY.
DR TIGRFAMS; TIGR00389; glyS_dimeric; 1.
DR PROSITE; PS50862; AA_TRNA_LIGASE-II; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.
SQ SEQUENCE 667 AA; 75410 MW; DD59506C0C9F18EF CRC64;

Query Match 39.0%; Score 44.5; DB 1; Length 667;
Best Local Similarity 36.7%; Pred. No. 42;
Matches 11; Conservative 4; Mismatches 6; Indels 9; Gaps 1;

QY 1 KTKRRKKQKV--KIA-----YEEIFVK 21
Db 145 KKKRKKVKQIKAVKLDVVVKEEILAK 174
| ||| :| :|:|:|:|

RESULT 13
SNCL_HUMAN STANDARD; PRT; 368 AA.
AC Q16533;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE snRNA activating protein complex 43 kDa subunit (SNAPc 43 kDa
DE subunit) (Proximal sequence element-binding transcription factor gamma
DE subunit) (PSE-binding factor gamma subunit) (PTF gamma subunit).
GN SNAPC1 OR SNAP43.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
RX MEDLINE=95231630; PubMed=7715707;
RA Henry R.W., Sadowski C.L., Kobayashi R., Hernandez N.;
RT "A TBP-TAF complex required for transcription of human snRNA genes by
RT RNA polymerase II and III.";
RL Nature 374:653-656(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96104548; PubMed=8524284;
RA Yoon J.B., Roeder R.G.;
RT "Cloning of two proximal sequence element-binding transcription factor
RT subunits (gamma and delta) that are required for transcription of
RT small nuclear RNA genes by RNA polymerases II and III and interact
RT with the TATA-binding protein.";
```

RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
RA Serrati L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
RA Cipriakovi G.V., Ussery D., Barrell B.G., Nurse P.;  
RT "The genome sequence of *Schizosaccharomyces pombe*.";  
RL Nature 415:871-880(2002).  
CC  
C -i- SIMILARITY: TO YEAST YNL091W.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC at the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement. (See <http://www.isb-sib.ch/announcement>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

|    |                               |           |
|----|-------------------------------|-----------|
| DR | ENBL: Z66525; CAA91432.1; -   |           |
| DR | ENBL: A4136538; CAB66473.1; - |           |
| KW | Hypothetical protein.         |           |
| FT | DOMAIN 33                     | 46        |
| FT | DOMAIN 184                    | 194       |
| FT | DOMAIN 695                    | 698       |
| FT | DOMAIN 714                    | 717       |
| FT | DOMAIN 718                    | 721       |
| FT | DOMAIN 938                    | 941       |
|    |                               | POLY-LYS. |
|    |                               | POLY-LYS. |
|    |                               | POLY-GLU. |
|    |                               | POLY-GLU. |
|    |                               | POLY-ARG. |
|    |                               | POLY-ARG. |

|    |                       |          |            |                  |              |
|----|-----------------------|----------|------------|------------------|--------------|
| SQ | SEQUENCE              | 1085 AA; | 122940 MW; | 041164132676F233 | CRC64;       |
|    | Query Match           | 37.7%    | Score 43;  | DB 1;            | Length 1085; |
|    | Best Local Similarity | 47.1%    | Pred No 1; | 2e+02;           |              |

[illegible]

| RESULT 15  |                       |                             |                         |          |  |
|------------|-----------------------|-----------------------------|-------------------------|----------|--|
| KPCI_NEUCR |                       |                             |                         |          |  |
| ID         | KPCI_NEUCR            | STANDARD;                   | PRT;                    | 1142 AA. |  |
| AC         | P87253;               |                             |                         |          |  |
| DT         | 15-JUL-1999           | (Rel. 38,                   | Created)                |          |  |
| DT'        | 15-JUL-1999           | (Rel. 38,                   | Last sequence update)   |          |  |
| DE         | 16-OCT-2001           | (Rel. 40,                   | Last annotation update) |          |  |
| DE         | Protein kinase C-like | (EC 2.7.1.-).               |                         |          |  |
| OS         | Neurospora crassa.    |                             |                         |          |  |
| OC         | Eukaryota; Fungi;     | Ascomycota; Pezizomycotina; | Sordariomycetes;        |          |  |
| OC         | Sordariales;          | Sordiariaceae;              | Neurospora.             |          |  |
| OX         | NCBI_TaxID=5141;      |                             |                         |          |  |
| RN         | [1]                   |                             |                         |          |  |
|            | SEQUENCE FROM N.A.    |                             |                         |          |  |
| PA         |                       |                             |                         |          |  |

```

RT "Molecular cloning of a protein kinase C homologue from Neurospora
RT crassa.";
RT Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
RL -! SIMILARITY: CONTAINS 2 ZINC-DEPENDENT PHORBOL-ESTER AND DAG
CC BINDING DOMAINS.
CC CC
CC -! SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC PKC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Y12002; CAA72731.1; -.
DR HSP; Q63450; 1A06.
DR InterPro; IPR000008; C2.
DR InterPro; IPR002219; DAG_PE-bind.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR000961; pkinase_C.
DR InterPro; IPR000861; REM repeat.

```

DR InterPro: IPR002290; Ser\_thr\_pkinase.  
DR Pfam: PF00069; pkinase; 1.  
DR Pfam: PF00130; DAG\_PE-bind; 2.  
DR Pfam: PF00433; pkinase\_C; 1.  
DR Pfam: PF02185; HRI; 2.  
DR ProDom: PD000001; Euk\_pkinase; 1.  
DR SMART: SM00109; C1; 2.  
DR SMART: SM00239; C2; 1.  
DR SMART: SM00074; HRI; 2.  
DR SMART: SM00133; S\_TK\_X; 1.  
DR SMART: SM00220; S\_TKc; 1.  
DR PROSITE: PS00479; DAG\_PE\_BIND\_DOM\_1; 2.  
DR PROSITE: PS00081; DAG\_PE\_BIND\_DOM\_2; 2.  
DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE: PS00011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE: PS00108; PROTEIN\_KINASE\_ST; 1.  
KW Transferase; Serine/threonine-protein kinase; ATP-binding; Zinc;  
KW Phorbol-ester binding; Repeat.  
FT DOMAIN 458 505 PHORBOL-ESTER AND DAG BINDING 1.  
FT DOMAIN 526 576 PHORBOL-ESTER AND DAG BINDING 2.  
FT DOMAIN 670 679 POLY-ALA.  
FT DOMAIN 714 717 POLY-GLN.  
FT DOMAIN 746 754 POLY-GLN.  
FT NP\_BIND 817 1076 PROTEIN KINASE.  
FT NP\_BIND 823 831 ATP (BY SIMILARITY).  
FT BINDING 846 846 ATP (BY SIMILARITY).  
FT ACT\_SITE 942 942 BY SIMILARITY.  
SQ SEQUENCE 1142 AA; 127897 MW; C599012DID1F5970 CRC64;

Query Match 37.7%; Score 43; DB 1; Length 1142;  
Best Local Similarity 31.8%; Pred. No. 1.2e+02;  
Matches 7; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 KTKRRKKQKVYIEIFVKV 22  
DB 835 ETKSRKLYAIKVLKKEFIEN 856

Search completed: December 2, 2002, 10:02:50  
Job time : 3.28856 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run On: December 2, 2002, 10:00:15 ; Search time 8.23881 Seconds  
(without alignments)  
575.215 Million cell updates/sec

Title: US-09-741-106-10  
Perfect score: 114  
Sequence: 1 KTKRRKKQKVIAEIEFVKNM 23

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_21:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phase:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvirus:\*
- 16: sp\_bacteriaph:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB ID     | Description          |
|------------|-------|-------------|--------|-----------|----------------------|
| 1          | 91    | 79.8        | 396    | 6 Q28874  | Q28874 canis famil   |
| 2          | 56    | 49.1        | 272    | 10 O64594 | O64594 arabidopsis   |
| 3          | 52    | 45.6        | 4981   | 5 O77372  | O77372 plasmodium    |
| 4          | 50    | 43.9        | 303    | 4 Q99566  | Q99566 homo sapien   |
| 5          | 50    | 43.9        | 304    | 4 Q9BPX2  | Q9BPX2 homo sapien   |
| 6          | 50    | 43.9        | 312    | 4 Q99794  | Q99794 homo sapien   |
| 7          | 50    | 43.9        | 312    | 4 P78549  | P78549 homo sapien   |
| 8          | 49    | 43.0        | 114    | 16 Q97G93 | Q97G93 clostridium   |
| 9          | 49    | 43.0        | 298    | 16 Q98R67 | Q98R67 mycoplasma    |
| 10         | 49    | 43.0        | 357    | 17 O29113 | O29113 archaetoglob  |
| 11         | 48    | 42.1        | 142    | 10 Q9FST8 | Q9FST8 gnetum gnetum |
| 12         | 48    | 42.1        | 231    | 10 Q9LZV2 | Q9LZV2 arabidopsis   |
| 13         | 48    | 42.1        | 414    | 2 Q9F982  | Q9F982 bacillus st   |
| 14         | 48    | 42.1        | 1941   | 4 Q9UEK4  | Q9UEK4 homo sapien   |
| 15         | 47.5  | 41.7        | 162    | 2 Q9REV3  | Q9REV3 salmonella    |
| 16         | 47.5  | 41.7        | 162    | 16 Q8ZRH4 | Q8ZRH4 salmonella    |

|    |      |      |      |           |                    |
|----|------|------|------|-----------|--------------------|
| 17 | 47.5 | 41.7 | 182  | 16 Q8Z921 | Q8Z921 salmonella  |
| 18 | 47   | 41.2 | 333  | 8 Q9BBQ4  | Q9BBQ4 lotus japon |
| 19 | 47   | 41.2 | 388  | 17 Q8TKE0 | Q8TKE0 methanosarc |
| 20 | 47   | 41.2 | 725  | 5 O9G101  | O9G101 trypanosoma |
| 21 | 47   | 41.2 | 725  | 5 Q9TZN8  | Q9TZN8 trypanosoma |
| 22 | 47   | 41.2 | 734  | 5 Q9VQK7  | Q9VQK7 drosophila  |
| 23 | 47   | 41.2 | 749  | 5 Q95RC1  | Q95RC1 drosophila  |
| 24 | 47   | 41.2 | 766  | 13 P79994 | P79994 gallus gall |
| 25 | 47   | 41.2 | 780  | 16 Q9WY48 | Q9WY48 thermotoga  |
| 26 | 47   | 41.2 | 1113 | 5 Q27481  | Q27481 caenorhabdi |
| 27 | 47   | 41.2 | 1599 | 11 Q99NH0 | Q99NH0 mus musculu |
| 28 | 47   | 41.2 | 1633 | 13 Q90941 | Q90941 gallus gall |
| 29 | 46.5 | 40.8 | 402  | 16 O48345 | O48345 anabaena sp |
| 30 | 46.5 | 40.8 | 413  | 16 O8VLT9 | O8VLT9 anabaena sp |
| 31 | 46.5 | 40.8 | 553  | 16 O87097 | O87097 aquifex aeo |
| 32 | 46   | 40.4 | 287  | 16 Q8R5P3 | Q8R5P3 thermoanaer |
| 33 | 45   | 39.5 | 84   | 12 Q91E08 | Q91E08 cydia pomon |
| 34 | 45   | 39.5 | 129  | 11 P70632 | P70632 rattus norv |
| 35 | 45   | 39.5 | 164  | 11 Q9CVB3 | Q9CVB3 mus musculu |
| 36 | 45   | 39.5 | 330  | 10 Q9LU74 | Q9LU74 arabidopsis |
| 37 | 45   | 39.5 | 532  | 8 O21282  | O21282 reclinomona |
| 38 | 45   | 39.5 | 571  | 4 Q9HA30  | Q9HA30 homo sapien |
| 39 | 45   | 39.5 | 671  | 4 Q96CL3  | Q96CL3 homo sapien |
| 40 | 45   | 39.5 | 702  | 16 Q8XI42 | Q8XI42 clostridium |
| 41 | 45   | 39.5 | 712  | 10 Q93WP2 | Q93WP2 chlamydomon |
| 42 | 45   | 39.5 | 764  | 4 Q8WUG7  | Q8WUG7 homo sapien |
| 43 | 45   | 39.5 | 791  | 4 Q8WXE1  | Q8WXE1 homo sapien |
| 44 | 45   | 39.5 | 1029 | 5 Q95TZ9  | Q95TZ9 drosophila  |
| 45 | 45   | 39.5 | 1410 | 5 O97230  | O97230 plasmodium  |

## ALIGNMENTS

## RESULT 1

Q28874 ID Q28874 PRELIMINARY; PRT; 396 AA.  
 AC Q28874;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE Tissue factor pathway inhibitor.  
 OS Canis familiaris (Dog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
 OX NCBI\_TaxID=9615;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95071310; PubMed=7980463;  
 RA Girard T.J., Gailani D., Broze G.J.Jr.;  
 RT "Complementary DNA sequencing of canine tissue factor pathway  
 RT inhibitor reveals a unique nanomeric repetitive sequence between the  
 RT second and third Kunitz domains.";  
 RL Biochem. J. 303:923-928(1994).  
 DR EMBL; S75369; AAB32443.1; -;  
 DR HSSP; P10646; ITEX.  
 DR InterPro; IPR002223; Kunitz\_BPTI.  
 DR Pfam; PF00014; Kunitz\_BPTI; 3.  
 DR PRINTS; PR00759; BASICPTASE.  
 DR ProDom; PD000222; Kunitz\_BPTI; 3.  
 DR SMART; SM00131; KU; 3.  
 DR PROSITE; PS00280; BPTI\_KUNITZ\_1; 3.  
 DR PROSITE; PS00279; BPTI\_KUNITZ\_2; 3.  
 KW Serine protease inhibitor.  
 SQ SEQUENCE 396 AA; 43948 MW; 50F65C8337A003D9 CRC64;

Query Match 79.8%; Score 91; DB 6; Length 396;

Best Local Similarity 78.3%; Pred. No. 1.8e-05;

Matches 18; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 KTKRRKKQKVIAEIEFVKNM 23

||||||| ||| ||| ||| :

DB 374 KTKRRKKQKVIAEIEFVKNM 396



```
RC TISSUE=LUNG;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG;
RA Strausberg R.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC003014; AAH03014.1; -.
DR EMBL; BC000391; AAH00391.1; -.
DR InterPro; IPR004036; EndoIII_Hhh.
DR InterPro; IPR003265; Endo_3c.
DR InterPro; IPR003651; Fes_bind.
DR Pfam; PF00730; HHH-GPD; 1.
DR SMART; SM00478; ENDO3c; 1.
DR SMART; SM00525; FES; 1.
DR PROSITE; PS01155; ENDONUCLEASE_III_2; 1.
KW Endonuclease.
SQ SEQUENCE 304 AA; 33569 MW; DA97D508BE3D83F0 CRC64;
Query Match 43.9%; Score 50; DB 4; Length 304;
Best Local Similarity 64.3%; Pred. No. 16;
Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
QY 3 KRRKKQKVKIAYE 16
Db 48 KRPRKAQRLRVAYE 61
RESULT 6
ID Q99794 PRELIMINARY; PRT; 312 AA.
AC Q99794;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-MAY-2002 (TrEMBLrel. 20, Last annotation update)
DE Endonuclease III homolog 1, hNTH1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97144402; PubMed=8990169;
RA Aspinwall R., Rothwell D.G., Roldan-Arjona T., Anselmino C.,
RA Ward C.J., Cheadle J.P., Sampson J.R., Lindahl T., Harris P.C.,
RA Hickson I.D.;
RT "Cloning and characterization of a functional human homolog of
RT Escherichia coli endonuclease III."
RL Proc. Natl. Acad. Sci. U.S.A. 94:109-114(1997).
DR EMBL; U79718; AAB41534.1; -.
DR InterPro; IPR004036; EndoIII_Hhh.
DR InterPro; IPR003265; Endo_3c.
DR InterPro; IPR003651; Fes_bind.
DR Pfam; PF00730; HHH-GPD; 1.
DR SMART; SM00478; ENDO3c; 1.
DR SMART; SM00525; FES; 1.
DR PROSITE; PS01155; ENDONUCLEASE_III_2; 1.
SQ SEQUENCE 312 AA; 34341 MW; B01821F89DDE8407 CRC64;
Query Match 43.9%; Score 50; DB 4; Length 312;
Best Local Similarity 64.3%; Pred. No. 16;
Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
QY 3 KRRKKQKVKIAYE 16
Db 56 KRPRKAQRLRVAYE 69
RESULT 7
ID P78549 PRELIMINARY; PRT; 312 AA.
AC P78549;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE hNTH1 (Endonuclease III homolog) (Nth endonuclease III-like 1) (E. coli).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Ricke D.O., Bruce D., Mundt M., Doggett N., Munk C., Saunders E.,
RA Robinson D., Jones M., Buckingham J., Chasteen L., Thompson S.,
RA Goodwin L., Bryant J., Tesmer J., Meincke L., Longmire J., White S.,
RA Ueng S., Tatum O., Campbell C., Fawcett J., Maltbie M., Misra M.,
RA Deaven L.;
RT "Sequencing of Human Chromosome 16p13.3";
RT Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Ricke D.O.;
RT "Large Scale Sequence Analysis and Annotation with the Sequence
RT Comparison Analysis (SCAN) System.";
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 9-312 FROM N.A.
RC TISSUE=BONE MARROW;
RX MEDLINE=98370989; PubMed=9705289;
RA Ikeda S., Biswas T., Roy R., Izumi T., Boldogh I., Kurosky A.,
RA Sarker A.H., Seki S., Mitra S.;
RT "Purification and characterization of human NTH1, a homolog of
RT Escherichia coli endonuclease III. Direct identification of Lys-212 as
RT the active nucleophilic residue.";
RL J. Biol. Chem. 273:21585-21593(1998).
RN [4]
RP SEQUENCE OF 9-312 FROM N.A.
RX MEDLINE=96180675; PubMed=8611553;
RA Hilbert T.P., Boorstein R.J., Kung H.C., Bolton P.H., Xing D.,
RA Cunningham R.P., Teebor G.W.;
RT "Purification of a mammalian homologue of Escherichia coli
RT endonuclease III: identification of a bovine pyrimidine hydrate-
RT thymine glycol DNase/AP lyase by irreversible cross linking to a
RT thymine glycol-containing oligonucleotide.";
RL Biochemistry 35:2505-2511(1996).
RN [5]
RP SEQUENCE OF 9-312 FROM N.A.
RC TISSUE=PLACENTA;
RA Imai K., Sarker A.H., Akiyama K., Ikeda S., Yao M., Tsutsui K.,
RA Shohmori T., Seki S.;
RT "Genomic structure and sequence of a human homologue (NTH1/NTH1) of
RT Escherichia coli endonuclease III with those of the adjacent parts of
RT TSC2 and SLC9A3R2 genes.";
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RA Rieder M.J., Braun A.C., Montoya M.A., Chung M.-W., Nguyen C.P.,
RA Nguyen D.A., Livingston R.J., Poel C.L., Robertson P.D.,
RA Schackwitz W.S., Sherwood J.K., Witrak L.A., Nickerson D.A.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC005600; AAC34209.1; -.
DR EMBL; AB001575; BAA19413.1; -.
DR EMBL; U81285; AAC51136.1; -.
DR EMBL; AB014460; BAA32895.1; -.
DR EMBL; AF498098; AAM11786.1; -.
DR InterPro; IPR004036; EndoIII_Hhh.
DR InterPro; IPR003265; Endo_3c.
DR InterPro; IPR003651; Fes_bind.
DR Pfam; PF00730; HHH-GPD; 1.
DR SMART; SM00478; ENDO3c; 1.
DR SMART; SM00525; FES; 1.
DR PROSITE; PS01155; ENDONUCLEASE_III_2; 1.
```

KW Endonuclease.  
SQ SEQUENCE 312 AA; 34389 MW; 379816A1E0B45050 CRC64;

Query Match 43.9%; Score 50; DB 4; Length 312;  
Best Local Similarity 64.3%; Pred. No. 16;  
Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 3 KRRKKRKQVKIAYE 16  
|| || ||::|||  
Db 56 KRPRKAQRLRAYE 69

RESULT 8  
Q97G93 PRELIMINARY; PRT; 114 AA.  
AC Q97G93;  
DT 01-OCT-2001 (TrEMBLrel. 18, Created)  
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)  
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)  
DE Uncharacterized protein, homolog of Spirochaeta aurantia  
DE (gi:152901).  
DE GN CAC2476.  
OS Clostridium acetobutylicum.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;  
OC Clostridiales; Clostridiaceae; Clostridium.  
OX NCBI\_TaxID=1488;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;  
RX MEDLINE=21359325; PubMed=1146286;  
RA Noelling J., Berton G., Omelchenko M.V., Makarova K.S., Zeng Q.,  
RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,  
RA Tatusov R.N., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,  
RA Bennett G.L., Koonin E.V., Smith D.R.;  
RA "Genome sequence and comparative analysis of the solvent-producing  
RT bacterium Clostridium acetobutylicum.";  
RL J. Bacteriol. 183:4823-4838(2001).  
DR EMBL; AE007746; AAK80430.1; -.  
KW Complete proteome.  
SQ SEQUENCE 114 AA; 13153 MW; A798B5CD369922DD CRC64;

Query Match 43.0%; Score 49; DB 16; Length 114;  
Best Local Similarity 53.3%; Pred. No. 8;  
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 1 KTKRRKKQVKIAY 15  
|:||||:|::|||  
Db 100 KSKRRKKAKTKVAY 114

RESULT 9  
Q98R67 PRELIMINARY; PRT; 298 AA.  
AC Q98R67;  
DT 01-OCT-2001 (TrEMBLrel. 18, Created)  
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)  
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)  
DE Heat shock protein GRPE (Activation of DNAK).  
DE MYPUL1430.  
GN OS Mycoplasma pulmonis.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;  
OC Mycoplasmataceae; Mycoplasma.  
OX NCBI\_TaxID=2107;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=UAB CTIP;  
RX MEDLINE=21267165; PubMed=11353084;  
RA Chambaud I., Heilig R., Ferris S., Barbe V., Samson D., Galisson F.,  
RA Moszer I., Dybvig K., Wroblewski H., Viari A., Rocha E.P.C.,  
RA Blanchard A.;  
RA "The complete genome sequence of the murine respiratory pathogen  
RT Mycoplasma pulmonis.";  
RL Nucleic Acids Res. 29:2145-2153(2001).

```
OS Gnetum gnemon.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Gnetales; Gnetales; Gnetales; Gnetales; Gnetales;
OX NCBI_TaxID=3382;
[1]
RN SEQUENCE FROM N.A.
RX MEDLINE=20472995; PubMed=11018150;
RA Becker A., Winter K.U., Mayer B., Saedler H., Theissen G.;
RT "MADS-Box Gene diversity in seed plants 300 million years ago.";
RL Mol. Biol. Evol. 17:1425-1434(2000).
DR EMBL; AJ251554; CAC13990.1; -
DR InterPro; IPR002487; TF_Kbox.
DR Pfam; PF01486; K-box; 1.
DR SMART; SM00432; MADS; 1.
DR PROSITE; PS50066; MADS_BOX_2; 1.
FT NON_TER
SQ SEQUENCE 142 AA; 16555 MW; 87045944F89CC5EF CRC64;

Query Match 42.1%; Score 48; DB 10; Length 142;
Best Local Similarity 47.4%; Pred. No. 15;
Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 2 TKRKRKKQKVIAEIEFVK 20
|||: |||||
Db 102 TKRVLNRKIKLYEIEHV 120

RESULT 12
Q9LVZ2 PRELIMINARY; PRT; 231 AA.
AC Q9LVZ2;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE Translation initiation factor eIF-2 beta chain-like protein.
GN T20L15.210.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eumossids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
[1]
RN SEQUENCE FROM N.A.
RA Bevan M., Terry N., Ardiles W., Buyschaert C., Dasseville R.,
RA De Clerck R., De Keyser A., Neyt P., Rouze P., Van Den Daele H.,
RA Villarroel R., Gielen J., Van Montagu M., Bancroft I., Mewes H.W.,
RA Rudd S., Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL162351; CAB82764.1; -
DR InterPro; IPR002735; eIF5_eIF2B.
DR Pfam; PF01873; eIF5_eIF2B; 1.
DR ProDom; PD004078; eIF5_eIF2B; 1.
DR Initiation factor.
KW Initiation factor.
SQ SEQUENCE 231 AA; 26547 MW; 36AEE3C61B1B5B72 CRC64;

Query Match 42.1%; Score 48; DB 10; Length 231;
Best Local Similarity 45.0%; Pred. No. 24;
Matches 9; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Qy 3 KKRKRKKQKVIAEIEFVK 22
|||: |||: ||
Db 32 KKRKRKKQKPLREDIFQFN 51

RESULT 13
Q9F982 PRELIMINARY; PRT; 414 AA.
ID Q9F982
```

---

```
Q9F982;
AC 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE P-aminobenzoate synthase (Fragment).
OS Bacillus stearothermophilus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Geobacillus.
OX NCBI_TaxID=1422;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=V;
RA Vasquez C., Pichuanes S., Saavedra C.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF198621; AAG28534.1; -
DR HSP; Q06128; IQDL.
DR InterPro; IPR000350; Chorismate_bind.
DR Pfam; PF00425; chorismate_bind; 1.
DR PRINTS; PR00095; ANTSNTHASEI.
DR ProDom; PD000779; Chorismate_bind; 1.
FT NON_TER
SQ SEQUENCE 414 AA; 48417 MW; 794714B7FD4E1561 CRC64;

Query Match 42.1%; Score 48; DB 2; Length 414;
Best Local Similarity 47.4%; Pred. No. 41;
Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 3 KKRKRKKQKVIAEIEFVK 21
|||: |||: ||
Db 4 KKRKRRTVSYAYRWFLQ 22

RESULT 14
Q9UEK4 PRELIMINARY; PRT; 1941 AA.
ID Q9UEK4;
DT 01-NAY-2000 (TREMBlrel. 13, Created)
DT 01-NAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE DCRRI protein (Fragment).
GN DCRRI.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primate; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN SEQUENCE FROM N.A.
RX MEDLINE=97396022; PubMed=9254009;
RA Eki T., Abe M., Naitou M., Sasanuma S., Nohata J., Kawashima K.,
RA Ahmad I., Hanaoka F., Murakami Y.;
RT "Cloning and characterization of novel gene, DCRRI, expressed from
RT Down's syndrome critical region of human chromosome 21q22.2.";
RL DNA Seq. 7:153-164(1997).
CC -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
DR EMBL; D83327; BAA23666.1; -
DR InterPro; IPR001440; TPR.
DR InterPro; IPR001841; Znf_ring.
DR Pfam; PF00515; TPR; 3.
DR Pfam; PF00097; zf-C3HC4; 1.
DR SMART; SM00184; RING; 1.
DR SMART; SM00028; TPR; 3.
KW Zinc-finger.
FT NON_TER
SQ SEQUENCE 1941 AA; 220196 MW; 1B9D07E926909C63 CRC64;

Query Match 42.1%; Score 48; DB 4; Length 1941;
Best Local Similarity 50.0%; Pred. No. 1.7e+02;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 KKRKRKKQKVIAEIEE 18
|||: |||: ||
Db 1093 RLKRRKRNKIKTKVEE 1110
```

RESULT 15  
Q9RFV3  
ID Q9RFV3 PRELIMINARY; PRT; 162 AA.  
AC Q9RFV3;  
DT 01-MAY-2000 (TRENBLrel. 13, Created)  
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)  
DT 01-MAY-2000 (TRENBLrel. 13, Last annotation update)  
DE Hypothetical 18.5 kDa protein.  
OS Salmonella typhimurium.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Salmonella.  
OX NCBI\_TaxID=602;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC14028;  
RX MEDLINE=20033552; PubMed=10564816;  
RA Altier C.; Suyemoto M.;  
RT "A recombinase-based selection of differentially expressed bacterial  
genes.";  
RL Gene 240:99-106(1999).  
DR EMBL; AF166275; AAF25608.1; -.  
KW Hypothetical protein.  
SQ SEQUENCE 162 AA; 18488 MW; ED814942A1106B12 CRC64;  
Query Match 41.7%; Score 47.5; DB 2; Length 162;  
Best Local Similarity 36.7%; Pred. No. 20;  
Matches 11; Conservative 5; Mismatches 7; Indels 7; Gaps 1;  
QY 1 KTKR-----KQKQKQKIAVEEIFVKNM 23  
Db 67 KMKRITLQYEIKTKDNGVKILYRDVYMKNL 96

Search completed: December 2, 2002, 10:04:10  
Job time : 11.2388 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run On: December 2, 2002, 09:59:15 ; Search time 7.61194 Seconds  
(without alignments)  
297.593 Million cell updates/sec

Title: US-09-741-106-12

Perfect score: 86

Sequence: 1 AKLNCRLRYKANKSSKL 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_101002.\*

```

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID       | Description        |
|------------|-------|-------------|--------|----------|--------------------|
| 1          | 86    | 100.0       | 17     | AAW06883 | Glycosaminoglycan  |
| 2          | 86    | 100.0       | 17     | AAW06883 | Anti-thrombin III  |
| 3          | 86    | 100.0       | 19     | ABE06704 | Bleeding tendency  |
| 4          | 86    | 100.0       | 19     | AAW48117 | Heparin-related pe |
| 5          | 86    | 100.0       | 432    | AAW59840 | Mature protein seq |
| 6          | 86    | 100.0       | 464    | AAW30445 | Sequence of human  |
| 7          | 86    | 100.0       | 464    | AAW10381 | Antithrombin III m |
| 8          | 86    | 100.0       | 464    | AAW10383 | Antithrombin III m |
| 9          | 86    | 100.0       | 464    | AAW10384 | Antithrombin III m |
| 10         | 86    | 100.0       | 464    | AAW10390 | Antithrombin III P |

|    |    |       |     |    |          |                    |
|----|----|-------|-----|----|----------|--------------------|
| 11 | 86 | 100.0 | 464 | 12 | AAW10380 | Antithrombin III m |
| 12 | 86 | 100.0 | 464 | 12 | AAW10382 | Antithrombin III m |
| 13 | 86 | 100.0 | 464 | 12 | AAW10385 | Antithrombin III m |
| 14 | 86 | 100.0 | 464 | 12 | AAW10386 | Antithrombin III m |
| 15 | 86 | 100.0 | 464 | 12 | AAW10387 | Antithrombin III m |
| 16 | 86 | 100.0 | 464 | 12 | AAW10388 | Antithrombin III m |
| 17 | 86 | 100.0 | 464 | 12 | AAW10389 | Antithrombin III m |
| 18 | 86 | 100.0 | 464 | 14 | AAW42895 | Human antithrombin |
| 19 | 86 | 100.0 | 464 | 14 | AAW42896 | Human antithrombin |
| 20 | 86 | 100.0 | 464 | 14 | AAW42897 | Human antithrombin |
| 21 | 86 | 100.0 | 464 | 14 | AAW42898 | Human antithrombin |
| 22 | 86 | 100.0 | 464 | 14 | AAW42899 | Human antithrombin |
| 23 | 86 | 100.0 | 464 | 14 | AAW42900 | Human antithrombin |
| 24 | 86 | 100.0 | 464 | 14 | AAW42901 | Human antithrombin |
| 25 | 86 | 100.0 | 464 | 14 | AAW42902 | Human antithrombin |
| 26 | 86 | 100.0 | 464 | 14 | AAW42903 | Human antithrombin |
| 27 | 86 | 100.0 | 464 | 14 | AAW42904 | Human antithrombin |
| 28 | 86 | 100.0 | 464 | 14 | AAW42905 | Human antithrombin |
| 29 | 86 | 100.0 | 464 | 14 | AAW42906 | Human antithrombin |
| 30 | 86 | 100.0 | 464 | 14 | AAW42907 | Human antithrombin |
| 31 | 86 | 100.0 | 464 | 14 | AAW42908 | Human antithrombin |
| 32 | 86 | 100.0 | 464 | 14 | AAW42909 | Human antithrombin |
| 33 | 86 | 100.0 | 464 | 14 | AAW42910 | Human antithrombin |
| 34 | 86 | 100.0 | 464 | 14 | AAW42911 | Human antithrombin |
| 35 | 86 | 100.0 | 464 | 14 | AAW42912 | Human antithrombin |
| 36 | 86 | 100.0 | 464 | 14 | AAW42913 | Human antithrombin |
| 37 | 86 | 100.0 | 464 | 14 | AAW42914 | Human antithrombin |
| 38 | 86 | 100.0 | 464 | 14 | AAW42915 | Human antithrombin |
| 39 | 86 | 100.0 | 464 | 14 | AAW42916 | Human antithrombin |
| 40 | 86 | 100.0 | 464 | 14 | AAW42917 | Human antithrombin |
| 41 | 86 | 100.0 | 464 | 14 | AAW42918 | Human antithrombin |
| 42 | 86 | 100.0 | 464 | 14 | AAW42933 | Human antithrombin |
| 43 | 86 | 100.0 | 464 | 14 | AAW42934 | Human antithrombin |
| 44 | 86 | 100.0 | 464 | 14 | AAW42935 | Human antithrombin |
| 45 | 86 | 100.0 | 464 | 21 | AAW92224 | Human antithrombin |

#### ALIGNMENTS

##### RESULT 1

AAW06883

ID AAW06883 standard; Peptide; 17 AA.

XX

AC AAW06883;

XX

DT 18-MAR-1997 (first entry)

XX

DE Glycosaminoglycan binding peptide from antithrombin III.

XX

KW Complement inhibitor; membrane co-factor protein; MCP;

KW decay accelerating factor; DAF; chimeric protein; glycosaminoglycan;

KW heparin; cell lysis; sepsis; adult respiratory distress syndrome;

KW reperfusion injury; cell damage; antithrombin III.

OS Synthetic.

XX

PN WO9634965-A2.

XX

PD 07-NOV-1996.

XX

PF 03-MAY-1996; 96WO-US06301.

XX

PR 05-MAY-1995; 95US-0435149.

XX

PA (CHIR ) CHIRON CORP.

XX

PI Creasey AA, Innis MA, Zaror I;

XX

WPI; 1996-506167/50.

XX

DR Chimeric proteins for inhibiting complement-mediated cell lysis -

XX

PT comprise membrane co-factor protein and decay accelerating factor

PT peptide sequences  
 XX Disclosure; Page 26; 33pp; English.  
 XX  
 CC A glycosaminoglycan binding peptide (AAW06883) derived from  
 CC antithrombin III is used in novel chimeric proteins of the formula  
 CC A-R1-B-R2-C, where A and C are peptides (AAW06875-79, AAW06883-90) able  
 CC to bind glycosaminoglycans (esp. heparin) present on cell surfaces,  
 CC R1 is membrane co-factor protein (MCP) or decay accelerating factor  
 CC (DAF), R2 is DAF when R1 is MCP or MCP when R1 is DAF, and B is a  
 CC peptide that may have complement inhibitor activity. The chimeric  
 CC proteins (see also AAW06882) are directed to cell surfaces where they  
 CC inhibit complement-mediated cell lysis. They are used to treat and  
 CC prevent disease states in which complement plays a role, e.g.  
 CC sepsis, adult respiratory distress syndrome, reperfusion injury and  
 CC tissue damage.  
 XX  
 SQ Sequence 17 AA;  
 Query Match 100.0%; Score 86; DB 17; Length 17;  
 Best Local Similarity 100.0%; Pred. No. 7.6e-08;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AKLNCRLYRKANKSSKL 17  
 IIIIIIIIIIIIIIIIIIII  
 Db 1 AKLNCRLYRKANKSSKL 17  
 IIIIIIIIIIIIIIIIIIII  
 RESULT 2  
 AAR92268  
 ID AAR92268 standard; peptide; 17 AA.  
 XX  
 AC AAR92268;  
 XX  
 DT 30-SEP-1996 (first entry)  
 XX  
 DE Anti-thrombin III heparin binding domain.  
 XX  
 KW Kunitz-type domain; C-terminal tail; tissue factor pathway inhibitor;  
 KW TFPI; TFPI-2; cell surface localisation; glycosaminoglycan; heparin;  
 KW phospholipid; binding; chimeric protein; mutin; substitution;  
 KW PI-reactive site; sepsis; septic shock; thrombosis; up-regulation;  
 KW tissue factor; injury; trauma; endotoxin; TNF; cancer; IL-1;  
 KW tumour necrosis factor; interleukin.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9604378-A2.  
 XX  
 PD 15-FEB-1996.  
 XX  
 PF 25-JUL-1995; 95WO-US09464.  
 XX  
 PR 05-AUG-1994; 94US-0286521.  
 XX  
 PA (CHIR ) CHIRON CORP.  
 XX  
 PI Creasey AA, Innis MA;  
 XX  
 XX WPI; 1996-129394/13.  
 DR  
 XX Chimeric protein comprising Kunitz-type domains from TFPI-1 and -2  
 PT - used for the treatment of septic shock and thrombosis disorders  
 PT  
 XX Claim 9; Page 11; 68pp; English.  
 PS  
 XX The sequences given in AAR92266-74 are heparin binding sites which were  
 CC used in the construction of the chimeric proteins or mutants of the  
 CC invention. These proteins comprise one of these heparin binding sites  
 CC in conjunction with a first, second and third Kunitz-type domain derived  
 CC from TFPI or TFPI-2. The Kunitz-type domain sequences are highly basic  
 CC and may be involved in cell surface localisation by glycosaminoglycan  
 CC (including heparin) or phospholipid binding. These mutants have one or

CC more substitutions exclusively in the PI-reactive site of one or more  
 CC Kunitz-type domains. The chimeric proteins and mutants may be used in a  
 CC pharmaceutical composition for the treatment of sepsis, septic shock and  
 CC thrombotic disorders. The proteins may be generally useful in the  
 CC treatment of diseases caused by the up-regulation of tissue factor  
 CC brought on by injury, trauma, endotoxin, TNF, cancer, IL-1 or other  
 CC agents or conditions.  
 XX  
 SQ Sequence 17 AA;  
 Query Match 100.0%; Score 86; DB 17; Length 17;  
 Best Local Similarity 100.0%; Pred. No. 7.6e-08;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AKLNCRLYRKANKSSKL 17  
 IIIIIIIIIIIIIIIIIIII  
 Db 1 AKLNCRLYRKANKSSKL 17  
 IIIIIIIIIIIIIIIIIIII  
 RESULT 3  
 ABB06704  
 ID ABB06704 standard; peptide; 19 AA.  
 XX  
 AC ABB06704;  
 XX  
 DT 10-JUN-2002 (first entry)  
 XX  
 DE Bleeding tendency treatment related peptide HBP-3 SEQ ID NO:14.  
 XX  
 KW Bleeding; inhibition; treatment; haemostatic; heparin;  
 KW extracorporeal circulation.  
 KW  
 XX Synthetic.  
 OS  
 XX Key Location/Qualifiers  
 FH Modified-site 19  
 FT /note= "amidated"  
 FT  
 XX JP2002003397-A.  
 PN  
 XX 09-JAN-2002.  
 PD  
 XX 23-JUN-2000; 2000JP-0189707.  
 PF  
 XX 23-JUN-2000; 2000JP-0189707.  
 PR  
 XX (ITOH-) ITO HAM KK.  
 PA  
 XX WPI; 2002-298299/34.  
 DR  
 XX Novel peptides useful for treatment of bleeding tendency -  
 PT Example 1; Page 7; 14pp; Japanese.  
 PS  
 XX The present invention describes peptides (I) used in the treatment of  
 CC bleeding tendency. (I) have haemostatic activity and can be used in the  
 CC inhibition of bleeding tendency. (I) can be used in the treatment of  
 CC bleeding tendency caused by the administration of heparin during  
 CC extracorporeal circulation. The present sequence represents a peptide  
 CC which is used in the exemplification of the present invention.  
 XX  
 SQ Sequence 19 AA;  
 Query Match 100.0%; Score 86; DB 23; Length 19;  
 Best Local Similarity 100.0%; Pred. No. 8.5e-08;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AKLNCRLYRKANKSSKL 17  
 IIIIIIIIIIIIIIIIIIII  
 Db 2 AKLNCRLYRKANKSSKL 18  
 IIIIIIIIIIIIIIIIIIII  
 RESULT 4



AAM48117  
 ID AAM48117 standard; peptide; 19 AA.  
 AC AAM48117;  
 XX  
 DT 04-MAR-2002 (first entry)  
 XX  
 DE Heparin-related peptide HBP-3.  
 XX  
 KW Heparin; antithrombotic.  
 XX  
 OS Synthetic.  
 XX  
 PN JP2001288198-A.  
 XX  
 PD 16-OCT-2001.  
 XX  
 PF 06-APR-2000; 2000JP-0104338.  
 XX  
 PR 06-APR-2000; 2000JP-0104338.  
 XX  
 PA (ITOH-) ITO HAM KK.  
 XX  
 DR WPI; 2002-093138/13.  
 XX  
 PT A heparin-highly affine peptide-combined carrier and separation and  
 PT purification of heparin antithrombotic active fraction -  
 PS  
 PS Claim 8; Page 6; 10pp; Japanese.  
 XX  
 CC The present invention relates to a peptide (AAM48117), which has high  
 CC affinity to heparin. Also claimed are (1) a carrier in which a peptide  
 CC (AAM48115) is combined; (2) a carrier in which a peptide (AAM48116/9)  
 CC having high affinity to heparin is combined; (3) a carrier in which a  
 CC peptide (AAM48118 and AAM48120) is combined; (4) a method for separating  
 CC and purifying an antithrombotic active fraction in heparin molecule in  
 CC which a solution containing heparin itself or a treated product of  
 CC heparin is contacted to the above carrier, and a heparin antithrombotic  
 CC active fraction prepared by the above method. The antithrombotic active  
 CC fraction in heparin molecule is useful as an antithrombotic agent. The  
 CC present peptide was used in the method of the invention.  
 XX  
 SQ Sequence 19 AA;  
 Query Match 100.0%; Score 86; DB 23; Length 19;  
 Best Local Similarity 100.0%; Pred. No. 8.5e-08;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AKLNCRLYRKANKSSKL 17  
 DB 2 AKLNCRLYRKANKSSKL 18  
 |||||  
 RESULT 5  
 AAW59840  
 ID AAW59840 standard; Protein; 432 AA.  
 XX  
 AC AAW59840;  
 XX  
 DT 20-NOV-1998 (first entry)  
 XX  
 DE Mature protein sequence of antithrombin III (ATIII).  
 XX  
 KW Protein expression; monocotyledon plant cell;  
 KW glycosylated alpha 1-antitrypsin; AAT; glycosylated antithrombin III;  
 KW ATIII; human serum albumin; HSA; subtilisin BPN'; treatment; emphysema;  
 KW antithrombotic; blood replacement.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W09836085-A1.  
 XX  
 PD 20-AUG-1998.

XX  
 PF 13-FEB-1998; 98WO-US03068.  
 XX  
 PR 13-FEB-1997; 97US-0038170.  
 PR 13-FEB-1997; 97US-0037991.  
 PR 13-FEB-1997; 97US-0038168.  
 PR 13-FEB-1997; 97US-0038169.  
 XX  
 PA (PHYT-) APPLIED PHYTOLOGICS INC.  
 XX  
 XX Rodriguez RL, Sutliff TD;  
 XX  
 DR WPI; 1998-467179/40.  
 DR N-PSDB; AAV41727.  
 XX  
 PT Expressing mature, glycosylated proteins in monocotyledonous plant  
 PT cells - from chimeric gene including signal peptide sequence,  
 PT specifically therapeutic agents and industrial enzymes  
 XX  
 PS Disclosure; Pages 29-30; 53pp; English.  
 XX  
 CC The present sequence represents the mature protein of antithrombin III  
 CC (ATIII). The protein is used to exemplify the invention. The  
 CC specification describes a method for producing mature heterologous  
 CC protein in monocotyledonous plant cells. The method comprises  
 CC transforming the cells with a chimeric gene comprising a monocotyledon  
 CC transcription regulator, inducible either during seed maturation or by  
 CC adding/removing a small molecule, DNA encoding the heterologous protein,  
 CC and DNA encoding a signal peptide, with the signal peptide causing  
 CC secretion of the protein from the cell. Proteins expressed in this  
 CC manner include mature glycosylated alpha 1-antitrypsin (AAT) with a  
 CC glycosylation pattern that significantly increases its serum half-life,  
 CC mature glycosylated antithrombin III (ATIII), mature human serum albumin  
 CC (HSA) having the native folding pattern as shown by bilirubin-binding  
 CC characteristics, or mature active subtilisin BPN'. These proteins are  
 CC useful therapeutically (e.g. AAT for treating emphysema, ATIII as  
 CC antithrombotic and HSA as blood replacement) or as industrial enzymes  
 CC (BPN' is used in detergents).  
 XX  
 SQ Sequence 432 AA;  
 Query Match 100.0%; Score 86; DB 19; Length 432;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-06;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AKLNCRLYRKANKSSKL 17  
 DB 124 AKLNCRLYRKANKSSKL 140  
 |||||  
 RESULT 6  
 AAP30445  
 ID AAP30445 standard; Protein; 464 AA.  
 XX  
 AC AAP30445;  
 XX  
 DT 25-MAY-1992 (first entry)  
 XX  
 DE Sequence of human antithrombin III (ATIII) from cDNA clones pa62 and  
 DE pa68.  
 XX  
 KW Thrombosis; therapy; cardiovascular disorders.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 1..32  
 FT /label= signal  
 XX  
 PN GB2116183-A.  
 XX  
 XX 21-SEP-1983.  
 PD



AAK10390  
ID AAK10390 standard: Protein: 464 AA.

```
FT Region 415..436
FT /label= P-region
FT /note= "substitutions occur in this region which is
FT positions 383-404 of mature ATIII"
XX
PN WO9100291-A.
XX
PD 10-JAN-1991.
XX
PF 23-JUN-1990; 90WO-EP01026.
XX
PR 26-JUN-1989; 89EP-0201675.
XX
PA (ALKU) AKZO NV.
XX
PI Dijkema R, Visser A;
XX
DR WPI; 1991-036710/05.
XX
PT New modified antithrombin III variants - with altered
PT heparin-dependent effect to factors Iia and Xa
XX
PS Claim 2; Fig 2; 24pp; English.
XX
CC Mutations introduced into the ATIII reactive centre P-region are
CC derived from the human prothrombin gene. Sequences known to be
CC physiological substrates for Factor Xa were selected and used to
CC replace the wild-type PvuII-StuI fragment of the original ATIII
CC cDNA sequence (see AAQ10323). (Numbers are for mature ATIII).
CC The substitutions are Ala(384) to Glu, Ser(385) to Leu,
CC Thr(386) to Leu, Ala(387) to Glu, Val(388) to Ser, Val(389) to
CC Tyr, Ala(391) to Asp, Ser(394) to Ile, Leu(395) to Val, Asn(396) to
CC Glu, Pro(397) to Gly, Asn(398) to Ser, Arg(399) to Asp, Val(400) to
CC Ala, Thr(401) to Glu, Phe(402) to Ile and Lys(403) to Gly. The
CC substitutions result in a shift in the heparin-dependent inhibition
CC profile of ATIII towards factors Iia and Xa.
CC See also AAR10381-RI0390.
XX
SQ Sequence 464 AA;
Query Match 100.0%; Score 86; DB 12; Length 464;
Best Local Similarity 100.0%; Pred. No. 2.3e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AKLNCRLYRKANKSSKL 17
Db 156 AKLNCRLYRKANKSSKL 172
RESULT 12
AAR10382
ID AAR10382 standard; Protein; 464 AA.
XX
AC AAR10382;
XX
DT 10-APR-1991 (first entry)
XX
DE Antithrombin III mutant #3.
XX
KW antithrombin 3; P-region variants; heparin-dependent; Factor Iia;
KW Factor Xa; meizothrombin; blood coagulation.
XX
FH Key Location/Qualifiers
FT Region 415..436
FT /label= P-region
FT /note= "substitutions occur in this region which is
FT positions 383-404 in mature ATIII"
XX
PN WO9100291-A.
XX
PD 10-JAN-1991.
XX
PF 23-JUN-1990; 90WO-EP01026.
XX
PR 26-JUN-1989; 89EP-0201675.
XX
PA (ALKU) AKZO NV.
XX
PI Dijkema R, Visser A;
XX
DR WPI; 1991-036710/05.
XX
PT New modified antithrombin III variants - with altered
PT heparin-dependent effect to factors Iia and Xa
XX
PS Claim 2; Fig 2; 24pp; English.
```

```
XX
PR 26-JUN-1989; 89EP-0201675.
XX
PA (ALKU) AKZO NV.
XX
PI Dijkema R, Visser A;
XX
DR WPI; 1991-036710/05.
XX
PT New modified antithrombin III variants - with altered
PT heparin-dependent effect to factors Iia and Xa
XX
PS Claim 2; Fig 2; 24pp; English.
XX
CC Mutations introduced into the ATIII reactive centre P-region are
CC derived from the human prothrombin gene. Sequences known to be
CC physiological substrates for Factor Xa were selected and used to
CC replace the wild-type PvuII-StuI fragment of the original ATIII
CC cDNA sequence (see AAQ10323). The substitutions are (given as
CC positions in mature ATIII): Ala(391) to Asp,
CC Ser(394) to Ile, Leu(395) to Val and Asn(396) to Glu. The
CC substitutions result in a shift in the heparin-dependent inhibition
CC profile of ATIII towards factors Iia and Xa.
CC See also AAR10380-1 and AAR10383-RI0390.
XX
SQ Sequence 464 AA;
Query Match 100.0%; Score 86; DB 12; Length 464;
Best Local Similarity 100.0%; Pred. No. 2.3e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AKLNCRLYRKANKSSKL 17
Db 156 AKLNCRLYRKANKSSKL 172
RESULT 13
AAR10385
ID AAR10385 standard; Protein; 464 AA.
XX
AC AAR10385;
XX
DT 10-APR-1991 (first entry)
XX
DE Antithrombin III mutant #6.
XX
KW antithrombin 3; P-region variants; heparin-dependent; Factor Iia;
KW Factor Xa; meizothrombin; blood coagulation.
XX
FH Key Location/Qualifiers
FT Region 415..436
FT /label= P-region
FT /note= "substitutions occur in this region which is
FT numbered 383-404 in mature ATIII"
XX
PN WO9100291-A.
XX
PD 10-JAN-1991.
XX
PF 23-JUN-1990; 90WO-EP01026.
XX
PR 26-JUN-1989; 89EP-0201675.
XX
PA (ALKU) AKZO NV.
XX
PI Dijkema R, Visser A;
XX
DR WPI; 1991-036710/05.
XX
PT New modified antithrombin III variants - with altered
PT heparin-dependent effect to factors Iia and Xa
XX
PS Claim 2; Fig 2; 24pp; English.
```

XX Mutations introduced into the ATIII reactive centre P-region are  
CC derived from the human prothrombin gene. Sequences known to be  
CC physiological substrates for Factor Xa were selected and used to  
CC replace the wild-type PvuII-StuI fragment of the original ATIII  
CC cDNA sequence (see AAQ10323). The substitutions (given as positions  
CC in mature ATIII) are Ser(394) to Ile and Asn(396) to Glu. The  
CC substitutions result in a shift in the heparin-dependent  
CC inhibition profile of ATIII towards factors IIA and Xa.  
CC See also AAR10380-4 and AAR10386-R10390.  
XX  
XX SQ Sequence 464 AA;

Query Match 100.0%; Score 86; DB 12; Length 464;  
Best Local Similarity 100.0%; Pred. No. 2.3e-06;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKLNCRLYRKANKSSKL 17  
Db 156 AKLNCRLYRKANKSSKL 172  
|||||

RESULT 14  
AAR10386  
ID AAR10386 standard; Protein; 464 AA.  
XX  
XX AC AAR10386;

XX 10-APR-1991 (first entry)  
XX  
XX DE Antithrombin III mutant #7.  
XX  
XX  
XX KW antithrombin 3; P-region variants; heparin-dependent; Factor IIA;  
XX Factor Xa; meizothrombin; blood coagulation.

XX  
XX FH Key Location/Qualifiers  
XX FT Region 415..436  
XX FT /label= P-region  
XX FT /note= "substitutions occur in this region which is  
XX positions 383-404 in mature ATIII"  
XX  
XX PN W09100291-A.  
XX  
XX PD 10-JAN-1991.

XX 23-JUN-1990; 90WO-EP01026.  
XX  
XX PF 26-JUN-1989; 89EP-0201675.  
XX  
XX PR (ALKU ) AKZO NV.

XX  
XX PA Dijkema R, Visser A;  
XX  
XX PI WPI; 1991-036710/05.  
XX  
XX DR  
XX  
XX PT New modified antithrombin III variants - with altered  
XX heparin-dependent effect to factors IIA and Xa

XX  
XX PS Claim 2; Fig 2; 24pp; English.  
XX  
XX CC Mutations introduced into the ATIII reactive centre P-region are  
XX derived from the human prothrombin gene. Sequences known to be  
XX physiological substrates for Factor Xa were selected and used to  
XX replace the wild-type PvuII-StuI fragment of the original ATIII  
XX cDNA sequence (see AAQ10323). The substitutions are Ala(391) to Asp  
XX and Asn(396) to Glu. (Numbers refer to mature ATIII). The  
XX substitutions result in a shift in the heparin-dependent inhibition  
XX profile of ATIII towards factors IIA and Xa.  
XX See also AAR10380-5 and AAR10387-R10390.

XX SQ Sequence 464 AA;  
Query Match 100.0%; Score 86; DB 12; Length 464;  
Best Local Similarity 100.0%; Pred. No. 2.3e-06;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKLNCRLYRKANKSSKL 17  
Db 156 AKLNCRLYRKANKSSKL 172  
|||||

Search completed: December 2, 2002, 10:02:24  
Job time : 8.61194 secs

Best Local Similarity 100.0%; Pred. No. 2.3e-06;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKLNCRLYRKANKSSKL 17  
Db 156 AKLNCRLYRKANKSSKL 172  
|||||

RESULT 15  
AAR10387  
ID AAR10387 standard; Protein; 464 AA.  
XX  
XX AC AAR10387;

XX 10-APR-1991 (first entry)  
XX  
XX DE Antithrombin III mutant #8.  
XX  
XX KW antithrombin 3; P-region variants; heparin-dependent; Factor IIA;  
XX Factor Xa; meizothrombin; blood coagulation.

XX  
XX FH Key Location/Qualifiers  
XX FT Region 415..436  
XX FT /label= P-region  
XX FT /note= "substitutions occur in this region which is  
XX positions 383-404 in mature ATIII"  
XX  
XX PN W09100291-A.  
XX  
XX PD 10-JAN-1991.

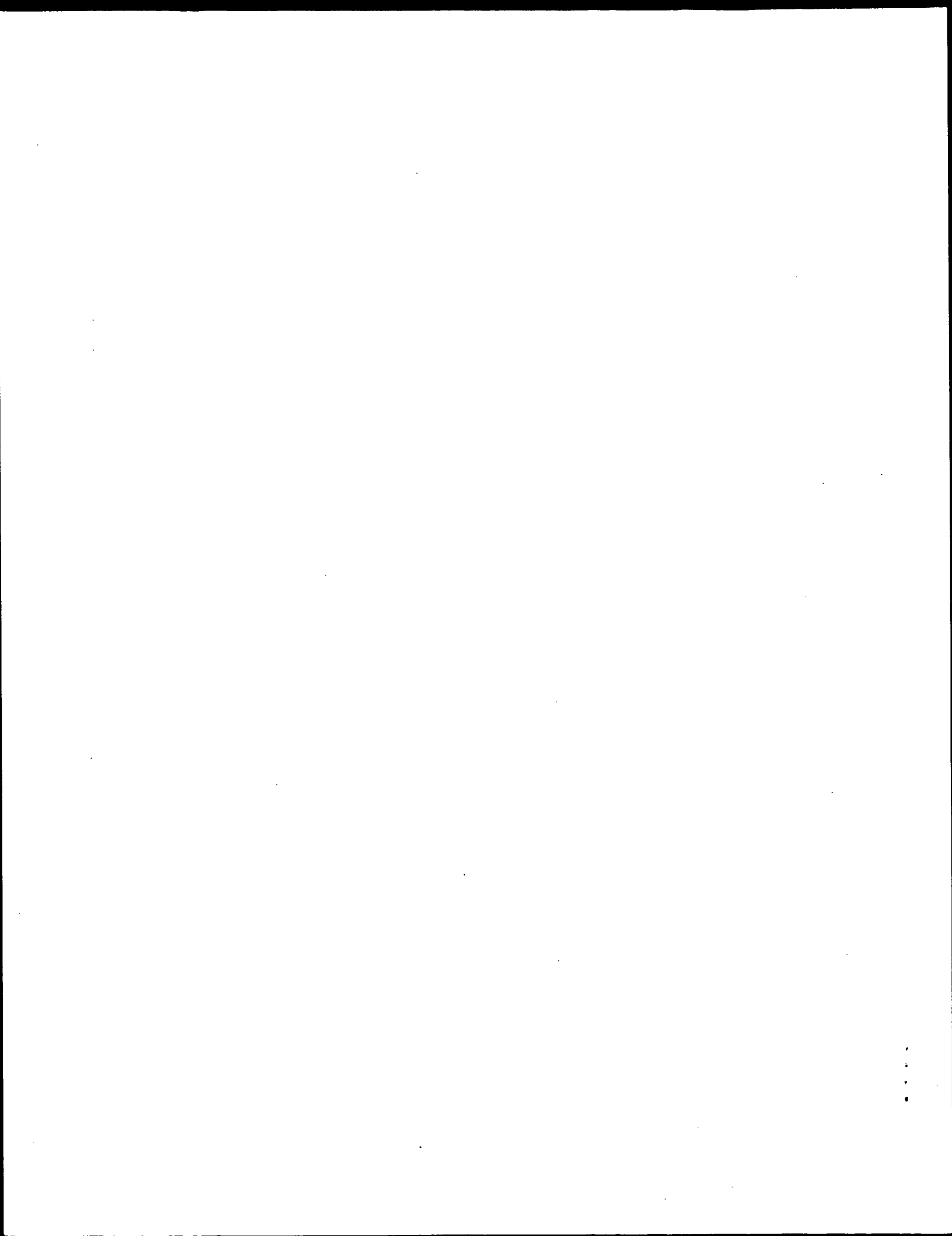
XX 23-JUN-1990; 90WO-EP01026.  
XX  
XX PR 26-JUN-1989; 89EP-0201675.  
XX  
XX PA (ALKU ) AKZO NV.

XX  
XX PI Dijkema R, Visser A;  
XX  
XX DR WPI; 1991-036710/05.  
XX  
XX PT New modified antithrombin III variants - with altered  
XX heparin-dependent effect to factors IIA and Xa

XX  
XX PS Claim 2; Fig 2; 24pp; English.  
XX  
XX CC Mutations introduced into the ATIII reactive centre P-region are  
XX derived from the human prothrombin gene. Sequences known to be  
XX physiological substrates for Factor Xa were selected and used to  
XX replace the wild-type PvuII-StuI fragment of the original ATIII  
XX cDNA sequence (see AAQ10323). The only substitution is Ala(391) to  
XX Asp. (Numbers refer to mature ATIII). The substitution results in  
XX a shift in the heparin-dependent inhibition profile of ATIII  
XX towards factors IIA and Xa.  
XX See also AAR10380-6 and AAR10388-R10390.

XX SQ Sequence 464 AA;  
Query Match 100.0%; Score 86; DB 12; Length 464;  
Best Local Similarity 100.0%; Pred. No. 2.3e-06;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKLNCRLYRKANKSSKL 17  
Db 156 AKLNCRLYRKANKSSKL 172  
|||||



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 2, 2002, 10:02:30 ; Search time 2.79104 Seconds  
(without alignments)  
179.212 Million cell updates/sec

Title: US-09-741-106-12

Perfect score: 86

Sequence: 1 AKLNCRLYRKANKSKL 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents\_AK:\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | % Match | Length | ID | Description         |
|------------|-------|---------|--------|----|---------------------|
| 1          | 86    | 100.0   | 17     | 1  | US-08-437-841-12    |
| 2          | 86    | 100.0   | 17     | 1  | US-08-286-521-12    |
| 3          | 86    | 100.0   | 17     | 1  | US-08-436-175-12    |
| 4          | 86    | 100.0   | 17     | 2  | US-08-435-149-8     |
| 5          | 86    | 100.0   | 17     | 4  | US-08-943-682-12    |
| 6          | 86    | 100.0   | 17     | 5  | PCT-US95-09464-12   |
| 7          | 86    | 100.0   | 464    | 1  | US-08-046-431A-2    |
| 8          | 80    | 93.0    | 465    | 3  | US-08-948-997-6     |
| 9          | 80    | 93.0    | 465    | 4  | US-09-348-817A-6    |
| 10         | 46    | 53.5    | 41     | 1  | US-08-612-986-8     |
| 11         | 46    | 53.5    | 41     | 1  | US-08-361-806A-8    |
| 12         | 46    | 53.5    | 41     | 5  | PCT-US95-16806A-8   |
| 13         | 43    | 50.0    | 20     | 1  | US-08-273-669-4     |
| 14         | 43    | 50.0    | 20     | 2  | US-08-954-724-4     |
| 15         | 43    | 50.0    | 20     | 2  | US-08-965-947-4     |
| 16         | 41    | 47.7    | 14     | 4  | US-09-675-922-10    |
| 17         | 41    | 47.7    | 20     | 1  | US-08-273-669-1     |
| 18         | 41    | 47.7    | 20     | 1  | US-08-273-669-6     |
| 19         | 41    | 47.7    | 20     | 2  | US-08-954-724-1     |
| 20         | 41    | 47.7    | 20     | 2  | US-08-954-724-6     |
| 21         | 41    | 47.7    | 20     | 2  | US-08-965-947-1     |
| 22         | 41    | 47.7    | 20     | 2  | US-08-965-947-6     |
| 23         | 41    | 47.7    | 22     | 4  | US-09-675-922-28    |
| 24         | 41    | 47.7    | 348    | 1  | US-08-118-270-13    |
| 25         | 41    | 47.7    | 348    | 5  | PCT-US93-08528-13   |
| 26         | 41    | 47.7    | 498    | 4  | US-09-238-303-8     |
| 27         | 40    | 46.5    | 322    | 4  | US-09-134-001C-3194 |

28 39 45.3 20 1 US-08-273-669-5 Sequence 5, Appli  
29 39 45.3 20 2 US-08-954-724-5 Sequence 5, Appli  
30 39 45.3 20 2 US-08-965-947-5 Sequence 5, Appli  
31 39 45.3 944 2 US-08-867-941-23 Sequence 23, Appli  
32 39 45.3 944 2 US-08-867-941-24 Sequence 24, Appli  
33 39 45.3 944 4 US-09-074-658-23 Sequence 23, Appli  
34 39 45.3 944 4 US-09-074-658-24 Sequence 24, Appli  
35 38 44.2 375 4 US-09-252-149B-33 Sequence 33, Appli  
36 38 44.2 681 4 US-08-760-615-4 Sequence 4, Appli  
37 38 44.2 681 4 US-08-760-615-6 Sequence 6, Appli  
38 37 43.0 20 1 US-08-273-669-2 Sequence 2, Appli  
39 37 43.0 20 1 US-08-273-669-3 Sequence 3, Appli  
40 37 43.0 20 2 US-08-954-724-2 Sequence 2, Appli  
41 37 43.0 20 2 US-08-954-724-3 Sequence 3, Appli  
42 37 43.0 20 2 US-08-965-947-2 Sequence 2, Appli  
43 37 43.0 20 2 US-08-965-947-3 Sequence 3, Appli  
44 37 43.0 172 4 US-09-071-035-442 Sequence 442, App  
45 37 43.0 360 4 US-09-286-691-25 Sequence 25, Appli

#### ALIGNMENTS

RESULT 1  
US-08-437-841-12  
; Sequence 12, Application US/08437841  
; Patent No. 5563123  
; GENERAL INFORMATION:  
; APPLICANT: Innis, Michael  
; APPLICANT: Creasey, Abia  
; TITLE OF INVENTION: Chimeric Proteins  
; NUMBER OF SEQUENCES: 37  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Chiron Corporation  
; STREET: 4560 Horton St.  
; CITY: Emeryville  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94608  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30B  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/437,841  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/286,521  
; FILING DATE: 05-AUG-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Savereide, Paul B.  
; REGISTRATION NUMBER: 36,914  
; REFERENCE/DOCKET NUMBER: 0990.001  
; TELEPHONE: 510-601-2585  
; TELEFAX: 510-655-3542  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 17 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-437-841-12

Query Match 100.0%; Score 86; DB 1; Length 17;  
Best Local Similarity 100.0%; Pred. No. 1.4e-07;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKLNCRLYRKANKSKL 17  
|||||

Db 1 AKLNCRLYRKANKSSKL 17

RESULT 2

US-08-286-521-12

Sequence 12, Application US/08286521

Patent No. 5589359

GENERAL INFORMATION:

APPLICANT: Innis, Michael

APPLICANT: Creasey, Abia

TITLE OF INVENTION: Chimeric Proteins

NUMBER OF SEQUENCES: 37

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Chiron Corporation

STREET: 4560 Horton St.

CITY: Emeryville

STATE: CA

COUNTRY: USA

ZIP: 94608

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30B

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/286,521

FILING DATE: 05-AUG-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Savereide, Paul B.

REGISTRATION NUMBER: 36,914

REFERENCE/DOCKET NUMBER: 0990.001

TELEPHONE: 510-601-2585

TELEFAX: 510-655-3542

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:

LENGTH: 17 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-286-521-12

Query Match 100.0%; Score 86; DB 1; Length 17;

Best Local Similarity 100.0%; Pred. No. 1.4e-07;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKLNCRLYRKANKSSKL 17

Db 1 AKLNCRLYRKANKSSKL 17

RESULT 4

US-08-435-149-8

Sequence 8, Application US/08435149

Patent No. 5866402

GENERAL INFORMATION:

APPLICANT: Innis, Michael A.

APPLICANT: Zaror, Isabel

APPLICANT: Creasey, Abia A.

TITLE OF INVENTION: CHIMERIC MCP AND DAF PROTEINS WITH CELL

TITLE OF INVENTION: SURFACE LOCALIZING DOMAIN

NUMBER OF SEQUENCES: 26

CORRESPONDENCE ADDRESS:

ADDRESSEE: CHIRON CORPORATION

STREET: INTELLECTUAL PROPERTY - R440, P.O. BOX 8097

CITY: EMERYVILLE

STATE: CALIFORNIA

COUNTRY: U.S.A.

ZIP: 94662-8097

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/435,149

FILING DATE: 05-MAY-1995

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: SAVEREIDE, PAUL B.

REGISTRATION NUMBER: 36,914

REFERENCE/DOCKET NUMBER: 0989.001

TELEPHONE: (510) 601-2585

TELEFAX: (510) 655-3542

TELEX: N/A

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 17 amino acids

TYPE: amino acid

STRANDEDNESS: single

QY 1 AKLNCRLYRKANKSSKL 17

Db 1 AKLNCRLYRKANKSSKL 17

RESULT 3

US-08-436-175-12

Sequence 12, Application US/08436175

Patent No. 5695088

GENERAL INFORMATION:

APPLICANT: Innis, Michael

APPLICANT: Creasey, Abia

TITLE OF INVENTION: Chimeric Proteins

NUMBER OF SEQUENCES: 37

CORRESPONDENCE ADDRESS:

ADDRESSEE: Chiron Corporation

STREET: 4560 Horton St.

CITY: Emeryville

STATE: CA

COUNTRY: USA

ZIP: 94608

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS



;  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-435-149-8

Query Match 100.0%; Score 86; DB 2; Length 17;  
Best Local Similarity 100.0%; Pred. No. 1.4e-07;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKLNCRLYRKANKSSKL 17  
| | | | | | | | | | | | | | | | | | | | |  
Db 1 AKLNCRLYRKANKSSKL 17

## RESULT 5

US-08-943-682-12  
; Sequence 12, Application US/08943682  
; Patent No. 6174721  
; GENERAL INFORMATION:  
; APPLICANT: Innis, Michael  
; TITLE OF INVENTION: Chimeric Proteins  
; NUMBER OF SEQUENCES: 37  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Chiron Corporation  
; STREET: 4560 Horton St.  
; CITY: Emeryville  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94608

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30B  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/943,682  
; FILING DATE: 03-OCT-1997  
; CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/438,184  
; FILING DATE: 09-MAY-1995  
; APPLICATION NUMBER: US 08/286,521  
; FILING DATE: 05-AUG-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Savereide, Paul B.  
; REGISTRATION NUMBER: 36,914  
; REFERENCE/DOCKET NUMBER: 0990.001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 510-601-2585  
; TELEFAX: 510-655-3542  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 17 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-943-682-12

Query Match 100.0%; Score 86; DB 4; Length 17;  
Best Local Similarity 100.0%; Pred. No. 1.4e-07;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKLNCRLYRKANKSSKL 17  
| | | | | | | | | | | | | | | | | | | | |  
Db 1 AKLNCRLYRKANKSSKL 17

## RESULT 6

PCT-US95-09464-12  
; Sequence 12, Application PC/TUS9509464  
; GENERAL INFORMATION:  
; APPLICANT: CHIRON CORPORATION

;  
; TITLE OF INVENTION: Chimeric Proteins  
; NUMBER OF SEQUENCES: 37  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Chiron Corporation  
; STREET: 4560 Horton St.  
; CITY: Emeryville  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94608  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30B  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/09464  
; FILING DATE: 25-JULY-1995  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Savereide, Paul B.  
; REGISTRATION NUMBER: 36,914  
; REFERENCE/DOCKET NUMBER: 0990.100  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 510-601-2585  
; TELEFAX: 510-655-3542  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 17 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
PCT-US95-09464-12

Query Match 100.0%; Score 86; DB 5; Length 17;  
Best Local Similarity 100.0%; Pred. No. 1.4e-07;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKLNCRLYRKANKSSKL 17  
| | | | | | | | | | | | | | | | | | | | |  
Db 1 AKLNCRLYRKANKSSKL 17

## RESULT 7

US-08-046-431A-2  
; Sequence 2, Application US/08046431A  
; Patent No. 5420252  
; GENERAL INFORMATION:  
; APPLICANT: KATO, HIROYUKI  
; APPLICANT: YOSHITAKE, SHINJI  
; APPLICANT: SUZUKI, SUGURA  
; APPLICANT: SUZUKI, NOBORU  
; APPLICANT: SETO, TOSHIO  
; APPLICANT: NAGAOKA, NAOKO  
; APPLICANT: MIZUI, YOSHIHARU  
; TITLE OF INVENTION: HUMAN ANTITHROMBIN III MUTANT  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS, P.C.  
; STREET: 2026 Rambling Road  
; CITY: Kalamazoo  
; STATE: Michigan  
; COUNTRY: USA  
; ZIP: 49008-1699  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage  
; COMPUTER: IBM PC/XT/AT Compatible  
; OPERATING SYSTEM: MS-DOS 5.0  
; SOFTWARE: WordPerfect 5.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/046,431A  
; FILING DATE:  
; CLASSIFICATION: 435

```
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP90488
; FILING DATE: 10-APR-1992
; APPLICATION NUMBER: JP31855
; FILING DATE: 22-FEB-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Terryence F. Chapman
; REGISTRATION NUMBER: 32549
; REFERENCE/DOCKET NUMBER: Furuya Case 1286
; TELEPHONE: (616) 381-1156
; TELEFAX: (616) 381-5465
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 464 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; DESCRIPTION: protein
; US-08-046-431A-2

Query Match 100.0%; Score 86; DB 1; Length 464;
Best Local Similarity 100.0%; Pred. No. 3.6e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKLNCRLRYRKANKSSKL 17
| | | | | | | | | | | | | | | | | | | |
Db 156 AKLNCRLRYRKANKSSKL 172

RESULT 8
US-08-948-997-6
; Sequence 6, Application US/08948997
; Patent No. 6008020
; GENERAL INFORMATION:
; APPLICANT: HASTINGS, GREGG
; APPLICANT: COLEMAN, TIM
; APPLICANT: LAWRENCE, DANIEL
; TITLE OF INVENTION: BRAIN-ASSOCIATED INHIBITOR OF
; TISSUE-TYPE PLASMINOGEN ACTIVATOR
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HUMAN GENOME SCIENCES, INC.
; STREET: 9410 KEY WEST AVENUE
; CITY: ROCKVILLE
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/948,997
; FILING DATE: Oct-10-97
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: A. ANDERS BROOKES
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PF336
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 465 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-948-997-6
```

```
Query Match 93.0%; Score 80; DB 3; Length 465;
Best Local Similarity 94.1%; Pred. No. 3.2e-05;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AKLNCRLRYRKANKSSKL 17
| | | | | | | | | | | | | | | | | | | |
Db 157 AKLNCRLRYRKANKSSDL 173

RESULT 9
US-09-348-817A-6
; Sequence 6, Application US/09348817A
; Patent No. 6191260
; GENERAL INFORMATION:
; APPLICANT: Hastings et al.
; TITLE OF INVENTION: Brain-Associated Inhibitor of Tissue-Type Plasminogen
; ACTIVATOR
; FILE REFERENCE: PF336D1
; CURRENT APPLICATION NUMBER: US/09/348,817A
; CURRENT FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: 08/948,997
; PRIOR FILING DATE: 1997-10-10
; PRIOR APPLICATION NUMBER: 60/028,117
; PRIOR FILING DATE: 1996-10-11
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Mus musculus
; US-09-348-817A-6

Query Match 93.0%; Score 80; DB 4; Length 465;
Best Local Similarity 94.1%; Pred. No. 3.2e-05;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AKLNCRLRYRKANKSSKL 17
| | | | | | | | | | | | | | | | | | | |
Db 157 AKLNCRLRYRKANKSSDL 173

RESULT 10
US-08-612-986-8
; Sequence 8, Application us/08612986
; Patent No. 5770384
; GENERAL INFORMATION:
; APPLICANT: Elliot J. Androphy
; APPLICANT: Dave E. Breiding
; TITLE OF INVENTION: E2 BINDING PROTEINS
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lahive & Cockfield
; STREET: 60 State Street, suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,986
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/361,806
; FILING DATE: 22 DEC 1994
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Paul L.
; REGISTRATION NUMBER: 35,965
```

; REFERENCE/DOCKET NUMBER: NEP-004DV  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 227-7400  
; TELEFAX: (617) 227-5941  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 41 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-612-986-8

Query Match 53.5%; Score 46; DB 1; Length 41;  
Best Local Similarity 72.7%; Pred. No. 0.74;  
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 LNCRLYRKANK 13  
Db 30 INCRHYRKKNK 40

## RESULT 11

US-08-361-806A-8  
; Sequence 8, Application US/08361806A  
; Patent No. 5792833  
; GENERAL INFORMATION:  
; APPLICANT: Eliot J. Androphy  
; APPLICANT: Dave E. Breiding  
; TITLE OF INVENTION: E2 BINDING PROTEINS  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Lahive & Cockfield  
; STREET: 60 State Street, suite 510  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109-1875  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/361.806A  
; FILING DATE: 22 DEC 1994  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Myers, Paul L.  
; REGISTRATION NUMBER: 35,965  
; REFERENCE/DOCKET NUMBER: NEP-004  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 227-7400  
; TELEFAX: (617) 227-5941  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 41 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-361-806A-8

Query Match 53.5%; Score 46; DB 1; Length 41;  
Best Local Similarity 72.7%; Pred. No. 0.74;  
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 LNCRLYRKANK 13  
Db 30 INCRHYRKKNK 40

## RESULT 12

PCT-US95-16806A-8  
; Sequence 8, Application PC/TUS9516806A

; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: E2 Binding Proteins  
; NUMBER OF SEQUENCES: 21  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII (text)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/16806A  
; FILING DATE: December 22, 1995  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/361,806  
; FILING DATE: 22-DEC-1994  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 41 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
PCT-US95-16806A-8

Query Match 53.5%; Score 46; DB 5; Length 41;  
Best Local Similarity 72.7%; Pred. No. 0.74;  
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 LNCRLYRKANK 13  
Db 30 INCRHYRKKNK 40

RESULT 13  
US-08-273-669-4  
; Sequence 4, Application US/08273669  
; Patent No. 5707798  
; GENERAL INFORMATION:  
; APPLICANT: Brann, Mark R.  
; TITLE OF INVENTION: Identification of Ligands by Selective  
; TITLE OF INVENTION: Amplification of Cells Transfected with  
; NUMBER OF SEQUENCES: 13  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: No. 5707798a No. 5707798disk of No. 5707798th America  
; STREET: 405 Lexington Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10174  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/273,669  
; FILING DATE: 12-JUL-94  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: AGRIS, Cheryl H.  
; REGISTRATION NUMBER: 34,086  
; REFERENCE/DOCKET NUMBER: 4011.200-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 867-0123  
; TELEFAX: (212) 878-9655  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein

## US-08-273-669-4

Query Match 50.0%; Score 43; DB 1; Length 20;  
Best Local Similarity 53.3%; Pred. No. 1.1;  
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 3 LNCRLYRKANKSSKL 17  
| | | | | | | | | |  
Db 1 LYCRIYRVAEKRTKV 15

## RESULT 14

US-08-954-724-4

; Sequence 4, Application US/08954724

; Patent No. 5912132

; GENERAL INFORMATION:

; APPLICANT: Brann, Mark R.

; TITLE OF INVENTION: Identification of Ligands by Selective

; TITLE OF INVENTION: Amplification of Cells Transfected with

; NUMBER OF SEQUENCES: 13

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: NO. 59121320 No. 5912132disk of No. 5912132th America

; STREET: 405 Lexington Avenue

; CITY: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10174

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/954,724

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/273,669

; FILING DATE: 12-JUL-94

; ATTORNEY/AGENT INFORMATION:

; NAME: AGRIS, Cheryl H.

; REGISTRATION NUMBER: 34,086

; REFERENCE/DOCKET NUMBER: 4011.200-US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 867-0123

; TELEFAX: (212) 878-9655

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 20 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-954-724-4

Query Match 50.0%; Score 43; DB 2; Length 20;  
Best Local Similarity 53.3%; Pred. No. 1.1;  
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 3 LNCRLYRKANKSSKL 17  
| | | | | | | | | |  
Db 1 LYCRIYRVAEKRTKV 15

## RESULT 15

US-08-965-947-4

; Sequence 4, Application US/08965947

; Patent No. 5955281

; GENERAL INFORMATION:

; APPLICANT: Brann, Mark R.

; TITLE OF INVENTION: Identification of Ligands by Selective

; TITLE OF INVENTION: Amplification of Cells Transfected with

; TITLE OF INVENTION: Receptors  
; NUMBER OF SEQUENCES: 13

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: No. 59552810 No. 5955281disk of No. 5955281th America

; STREET: 405 Lexington Avenue

; CITY: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10174

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/965,947

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/273,669

; FILING DATE: 12-JUL-94

; ATTORNEY/AGENT INFORMATION:

; NAME: AGRIS, Cheryl H.

; REGISTRATION NUMBER: 34,086

; REFERENCE/DOCKET NUMBER: 4011.200-US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 867-0123

; TELEFAX: (212) 878-9655

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 20 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-965-947-4

Query Match 50.0%; Score 43; DB 2; Length 20;

Best Local Similarity 53.3%; Pred. No. 1.1;

Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 3 LNCRLYRKANKSSKL 17

| | | | | | | | | |

Db 1 LYCRIYRVAEKRTKV 15

Search completed: December 2, 2002, 10:05:58

Job time : 3.79104 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 2, 2002, 10:00:50 ; Search time 1.60697 Seconds  
(without alignments)  
168.461 Million cell updates/sec

Title: US-09-741-106-12

Perfect score: 86

Sequence: 1 AKLNCRLRYRKANKSSKL 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 102317 seqs, 15924203 residues

Total number of hits satisfying chosen parameters: 102317

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

#### Database :

Published\_Applications\_AA:\*  
1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description         |
|------------|-------|-------------|--------|-------|---------------------|
| 1          | 86    | 100.0       | 464    | 10    | US-09-414-834-1     |
| 2          | 80    | 93.0        | 465    | 9     | US-09-987-021-6     |
| 3          | 80    | 93.0        | 465    | 10    | US-09-957-485-6     |
| 4          | 41    | 47.7        | 14     | 10    | US-09-848-664-1     |
| 5          | 41    | 47.7        | 498    | 10    | US-09-946-239-8     |
| 6          | 38    | 44.2        | 375    | 9     | US-09-841-730-16    |
| 7          | 38    | 44.2        | 375    | 9     | US-09-859-211-31    |
| 8          | 37    | 43.0        | 221    | 9     | US-10-108-605-297   |
| 9          | 36.5  | 42.4        | 235    | 10    | US-09-864-761-43113 |
| 10         | 36.5  | 42.4        | 475    | 12    | US-10-142-373-2     |
| 11         | 36.5  | 42.4        | 478    | 10    | US-09-765-111A-27   |
| 12         | 36.5  | 42.4        | 505    | 10    | US-09-765-111A-16   |
| 13         | 36.5  | 42.4        | 506    | 12    | US-10-109-886-6     |
| 14         | 36.5  | 42.4        | 777    | 10    | US-09-765-111A-2    |
| 15         | 36.5  | 42.4        | 811    | 10    | US-09-765-111A-23   |
| 16         | 36.5  | 42.4        | 840    | 10    | US-09-765-111A-4    |
| 17         | 36.5  | 42.4        | 874    | 10    | US-09-765-111A-6    |
| 18         | 35    | 40.7        | 14     | 10    | US-09-848-664-4     |
| 19         | 35    | 40.7        | 136    | 10    | US-09-867-550-396   |

|    |    |      |      |    |                     |                   |
|----|----|------|------|----|---------------------|-------------------|
| 20 | 35 | 40.7 | 152  | 10 | US-09-815-242-10715 | Sequence 10715, A |
| 21 | 35 | 40.7 | 173  | 10 | US-09-393-634-23    | Sequence 23, Appl |
| 22 | 35 | 40.7 | 207  | 9  | US-09-976-736-13    | Sequence 13, Appl |
| 23 | 35 | 40.7 | 259  | 9  | US-09-976-736-11    | Sequence 11, Appl |
| 24 | 35 | 40.7 | 259  | 9  | US-09-976-736-12    | Sequence 12, Appl |
| 25 | 35 | 40.7 | 259  | 12 | US-10-052-586-250   | Sequence 250, App |
| 26 | 35 | 40.7 | 262  | 10 | US-09-764-864-978   | Sequence 978, App |
| 27 | 35 | 40.7 | 264  | 10 | US-09-741-669-374   | Sequence 374, App |
| 28 | 35 | 40.7 | 266  | 10 | US-09-954-197-2     | Sequence 2, Appl  |
| 29 | 35 | 40.7 | 280  | 10 | US-09-323-998D-37   | Sequence 37, Appl |
| 30 | 35 | 40.7 | 310  | 10 | US-09-393-634-7     | Sequence 7, Appl  |
| 31 | 35 | 40.7 | 344  | 10 | US-09-764-864-1413  | Sequence 1413, Ap |
| 32 | 35 | 40.7 | 739  | 10 | US-09-957-635-2     | Sequence 2, Appl  |
| 33 | 35 | 40.7 | 859  | 10 | US-09-529-063-70    | Sequence 70, Appl |
| 34 | 35 | 40.7 | 1686 | 12 | US-10-092-219-2     | Sequence 2, Appl  |
| 35 | 34 | 39.5 | 34   | 10 | US-09-864-761-39376 | Sequence 39376, A |
| 36 | 34 | 39.5 | 49   | 10 | US-09-764-887-204   | Sequence 204, App |
| 37 | 34 | 39.5 | 200  | 10 | US-09-815-242-5175  | Sequence 5175, Ap |
| 38 | 34 | 39.5 | 205  | 10 | US-09-206-639-5     | Sequence 5, Appl  |
| 39 | 34 | 39.5 | 229  | 10 | US-09-323-998D-38   | Sequence 38, Appl |
| 40 | 34 | 39.5 | 254  | 10 | US-09-954-737-8     | Sequence 8, Appl  |
| 41 | 34 | 39.5 | 281  | 9  | US-09-756-854-3     | Sequence 3, Appl  |
| 42 | 34 | 39.5 | 281  | 9  | US-10-041-574-3     | Sequence 3, Appl  |
| 43 | 34 | 39.5 | 335  | 10 | US-09-826-212-7     | Sequence 7, Appl  |
| 44 | 34 | 39.5 | 335  | 10 | US-09-802-669-2     | Sequence 2, Appl  |
| 45 | 34 | 39.5 | 335  | 10 | US-09-333-966-6     | Sequence 6, Appl  |

#### ALIGNMENTS

RESULT 1  
US-09-414-834-1  
; Sequence 1, Application US/09414834  
; Patent No. US20020076413A1  
; GENERAL INFORMATION:  
; APPLICANT: O'Reilly, Michael S.  
; APPLICANT: Folkman, M. Judah  
; APPLICANT: Pirie-Shepherd, Steven  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INHIBITING ANGIOGENESIS  
; FILE REFERENCE: 07555.0001  
; CURRENT APPLICATION NUMBER: US/09/414,834  
; CURRENT FILING DATE: 1999-10-08  
; NUMBER OF SEQ ID NOS: 1  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 464  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: Antithrombin III  
US-09-414-834-1

Query Match 100.0%; Score 86; DB 10; Length 464;  
Best Local Similarity 100.0%; Pred. No. 7.9e-07;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKLNCRLRYRKANKSSKL 17  
| | | | | | | | | | | | | | | | | | | | | |  
Db 156 AKLNCRLRYRKANKSSKL 172

RESULT 2  
US-09-987-021-6  
; Sequence 6, Application US/09987021  
; Patent No. US20020165147A1  
; GENERAL INFORMATION:  
; APPLICANT: Yepes, et al.  
; TITLE OF INVENTION: Brain-Associated Inhibitor of Tissue-Type Plasminogen Activator  
; FILE REFERENCE: PF336P2  
; CURRENT APPLICATION NUMBER: US/09/987,021  
; CURRENT FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: 09/957,485

; FILE REFERENCE: 2001-09-21  
; PRIOR FILING DATE: 2001-09-21  
; CURRENT APPLICATION NUMBER: 09/722,292  
; PRIOR FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: 60/247,971  
; PRIOR FILING DATE: 2000-11-14  
; PRIOR APPLICATION NUMBER: 09/521,664  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 09/348,817  
; PRIOR FILING DATE: 1999-07-08  
; PRIOR APPLICATION NUMBER: 60/123,704  
; PRIOR FILING DATE: 1999-03-10  
; PRIOR APPLICATION NUMBER: 08/948,997  
; PRIOR FILING DATE: 1997-10-10  
; PRIOR APPLICATION NUMBER: 60/028,117  
; PRIOR FILING DATE: 1996-10-11  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 465  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-987-021-6

Query Match 93.0%; Score 80; DB 9; Length 465;  
Best Local Similarity 94.1%; Pred. No. 7.5e-06;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AKLNCRLYRKANKSSKL 17  
|||||  
Db 157 AKLNCRLYRKANKSSDL 173

RESULT 3  
US-09-957-485-6  
; Sequence 6, Application US/09957485  
; Patent No. US20020143165A1  
; GENERAL INFORMATION:  
; TITLE OF INVENTION: Human Genome Sciences, Inc. et al.  
; APPLICANT: Human Genome Sciences, Inc. et al.  
; TITLE OF INVENTION: Brain-Associated Inhibitor of Tissue-Type Plasminogen  
; FILE REFERENCE: PF336P1  
; CURRENT APPLICATION NUMBER: US/09/957,485  
; CURRENT FILING DATE: 2001-09-21  
; PRIOR APPLICATION NUMBER: US 09/521,664  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: US 60/123,704  
; PRIOR FILING DATE: 1999-03-10  
; NUMBER OF SEQ ID NOS: 21  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 465  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-957-485-6

Query Match 93.0%; Score 80; DB 10; Length 465;  
Best Local Similarity 94.1%; Pred. No. 7.5e-06;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AKLNCRLYRKANKSSKL 17  
|||||  
Db 157 AKLNCRLYRKANKSSDL 173

RESULT 4  
US-09-848-664-1  
; Sequence 1, Application US/09848664  
; Patent No. US20020146414A1  
; GENERAL INFORMATION:  
; APPLICANT: Sakiyama-Elbert, Shelly E.  
; APPLICANT: Hubbell, Jeffrey A.  
; TITLE OF INVENTION: Controlled Release of No. US20020146414A1-Heparin Binding Growth  
; TITLE OF INVENTION: Factors from Heparin Containing Matrices

; FILE REFERENCE: ETH 108  
; CURRENT APPLICATION NUMBER: US/09/848,664  
; CURRENT FILING DATE: 2001-05-03  
; PRIOR APPLICATION NUMBER: 09/298,084  
; PRIOR FILING DATE: 1999-04-22  
; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 14  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: MOD\_RES  
; LOCATION: (2)  
; OTHER INFORMATION: Xaa is bala (Beta Alanine)  
US-09-848-664-1

Query Match 47.7%; Score 41; DB 10; Length 14;  
Best Local Similarity 81.8%; Pred. No. 0.47;  
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AKLNCRLYRKA 11  
|||  
Db 4 AKLAARDLYRKA 14

RESULT 5  
US-09-946-239-8  
; Sequence 8, Application US/09946239  
; Patent No. US20020044945A1  
; GENERAL INFORMATION:  
; APPLICANT: Barr, Margaret C.  
; TITLE OF INVENTION: No. US20020044945A1 Feline Immunodeficiency Virus Nucleotide  
; FILE REFERENCE: 18617.0059  
; CURRENT APPLICATION NUMBER: US/09/946,239  
; CURRENT FILING DATE: 2001-09-04  
; PRIOR APPLICATION NUMBER: US 09/238,303, US 60/072,927  
; PRIOR FILING DATE: 1999-01-28, 1998-01-29  
; NUMBER OF SEQ ID NOS: 17  
; SEQ ID NO 8  
; LENGTH: 498  
; TYPE: PRT  
; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: protein encoded by the gag gene of a recombinant viral  
; OTHER INFORMATION: clone constructed from the genomic DNA of a Pallas's cat felin  
; OTHER INFORMATION: immunodeficiency virus  
US-09-946-239-8

Query Match 47.7%; Score 41; DB 10; Length 498;  
Best Local Similarity 61.5%; Pred. No. 19;  
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 NCRLYRKANKSSK 16  
|||  
Db 423 NCRAPRKCKGK 435

RESULT 6  
US-09-841-730-16  
; Sequence 16, Application US/09841730  
; Patent No. US20020157126A1  
; GENERAL INFORMATION:  
; APPLICANT: Lee, Se-Jin  
; APPLICANT: McPherron, Alexandra C.  
; TITLE OF INVENTION: GROWTH DIFFERENTIATION FACTOR RECEPTORS,  
; TITLE OF INVENTION: AGONISTS AND ANTAGONISTS THEREOF, AND METHODS OF USING SAME  
; FILE REFERENCE: JHU1470-2  
; CURRENT APPLICATION NUMBER: US/09/841,730  
; CURRENT FILING DATE: 2001-04-24  
; PRIOR APPLICATION NUMBER: 09/626,896  
; PRIOR FILING DATE: 2000-07-27

; PRIOR APPLICATION NUMBER: 09/485,046  
; PRIOR FILING DATE: 2000-01-31  
; PRIOR APPLICATION NUMBER: PCT/US98/15598  
; PRIOR FILING DATE: 1998-07-28  
; PRIOR APPLICATION NUMBER: 60/054,461  
; PRIOR FILING DATE: 1997-08-01  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 16  
; LENGTH: 375  
; TYPE: PRT  
; ORGANISM: Ovine  
US-09-841-730-16

Query Match 44.2%; Score 38; DB 9; Length 375;  
Best Local Similarity 57.1%; Pred. No. 43;  
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 4 NCRLRKANKSSKL 17  
Db 40 NACLWRQNNKSSRL 53

RESULT 7  
US-09-859-211-31  
; Sequence 31, Application US/09859211  
; Patent No. US20020157125A1  
; GENERAL INFORMATION:  
; APPLICANT: Lee, Se-jin  
; APPLICANT: McPherron, Alexandra C.  
; TITLE OF INVENTION: GROWTH DIFFERENTIATION FACTOR-8  
; FILE REFERENCE: 07265/144001  
; CURRENT APPLICATION NUMBER: US/09/859,211  
; CURRENT FILING DATE: 2001-05-15  
; PRIOR APPLICATION NUMBER: 09/019,070  
; PRIOR FILING DATE: 1998-02-05  
; PRIOR APPLICATION NUMBER: 08/962,445  
; PRIOR FILING DATE: 1997-05-23  
; PRIOR APPLICATION NUMBER: 08/847,910  
; PRIOR FILING DATE: 1997-04-28  
; PRIOR APPLICATION NUMBER: 08/795,071  
; PRIOR FILING DATE: 1997-02-05  
; PRIOR APPLICATION NUMBER: 08/525,596  
; PRIOR FILING DATE: 1995-10-26  
; PRIOR APPLICATION NUMBER: PCT/US94/03019  
; PRIOR FILING DATE: 1994-03-18  
; PRIOR APPLICATION NUMBER: 08/033,923  
; PRIOR FILING DATE: 1993-03-19  
; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 31  
; LENGTH: 375  
; TYPE: PRT  
; ORGANISM: Ovine  
US-09-859-211-31

Query Match 44.2%; Score 38; DB 9; Length 375;  
Best Local Similarity 57.1%; Pred. No. 43;  
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 4 NCRLRKANKSSKL 17  
Db 40 NACLWRQNNKSSRL 53

RESULT 8  
US-10-108-605-297  
; Sequence 297, Application US/10108605  
; Patent No. US20020160934A1  
; GENERAL INFORMATION:  
; APPLICANT: Broadus, Julie  
; APPLICANT: Stam, Lynn  
; APPLICANT: Bachmann, Jane

; APPLICANT: Kamdar, Kim  
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES FROM DROSOPHILA MELANOGASTER THAT ENCOD  
; TITLE OF INVENTION: PROTEINS ESSENTIAL FOR LARVAL VIABILITY AND USES THEREOF  
; FILE REFERENCE: 31133B  
; CURRENT APPLICATION NUMBER: US/10/108,605  
; CURRENT FILING DATE: 2002-03-27  
; PRIOR APPLICATION NUMBER: US 09/761,142  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/176,418  
; PRIOR FILING DATE: 2000-01-14  
; NUMBER OF SEQ ID NOS: 361  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 297  
; LENGTH: 221  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
US-10-108-605-297

Query Match 43.0%; Score 37; DB 9; Length 221;  
Best Local Similarity 63.6%; Pred. No. 36;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 6 RLYRKANKSSK 16  
Db 61 RLYRKKNDAK 71

RESULT 9  
US-09-864-761-43113  
; Sequence 43113, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO  
; FILE REFERENCE: Acomica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21

```
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 43113
; LENGTH: 235
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC004613.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EST_HUMAN HIT: AI984820.1, EVALUE 1.00e-58
; OTHER INFORMATION: SWISSPROT HIT: Q61200, EVALUE 1.00e-129
US-09-864-761-43113

Query Match 42.4%; Score 36.5; DB 10; Length 235;
Best Local Similarity 50.0%; Pred. No. 47;
Matches 7; Conservative 4; Mismatches 2; Indels 1; Gaps 1;

QY 4 NCRL-YRKANKSSK 16
 |||:|:|:|:|
Db 169 NCRIEYKVDKATK 182

RESULT 10
US-10-142-373-2
; Sequence 2, Application US/10142373
; Patent No. US20020137665A1
; GENERAL INFORMATION:
; APPLICANT: EVANS, Ronald
; APPLICANT: FOWMAN, Barry
; TITLE OF INVENTION: MODULATORS OF PEROXISOME PROLIFERATOR ACTIVATED
; TITLE OF INVENTION: RECEPTOR-GAMMA,
; TITLE OF INVENTION: AND METHODS FOR THE USE THEREOF
; FILE REFERENCE: SAL1480-2
; CURRENT APPLICATION NUMBER: US/10/142,373
; CURRENT FILING DATE: 2002-05-08
; PRIOR APPLICATION NUMBER: US/09/788,070
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: US 09/955,302
; PRIOR FILING DATE: 1999-02-22
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 475
; TYPE: PRT
; ORGANISM: Mus Musculus
US-10-142-373-2

Query Match 42.4%; Score 36.5; DB 12; Length 475;
Best Local Similarity 58.3%; Pred. No. 97;
Matches 7; Conservative 4; Mismatches 0; Indels 1; Gaps 1;

QY 3 LNCRLYRKA-NK 13
 |||:|:|:|:|
Db 148 LNCRIHKSRNK 159

RESULT 11
US-09-765-111A-27
; Sequence 27, Application US/09765111A
; Patent No. US20020106796A1
; GENERAL INFORMATION:
; APPLICANT: Fletcher, Jonathan A.
; APPLICANT: Kroll, Todd G.
; TITLE OF INVENTION: PAX8-PPARGamma NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: AND POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: B0801/7196/ERP/MAT
; CURRENT APPLICATION NUMBER: US/09/765,111A
; CURRENT FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: US 60/177,109
; PRIOR FILING DATE: 2000-01-20
; PRIOR APPLICATION NUMBER: US 60/225,079
; PRIOR FILING DATE: 2000-08-14
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 27
; LENGTH: 478
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-765-111A-27

Query Match 42.4%; Score 36.5; DB 10; Length 478;
Best Local Similarity 58.3%; Pred. No. 98;
Matches 7; Conservative 4; Mismatches 0; Indels 1; Gaps 1;

QY 3 LNCRLYRKA-NK 13
 |||:|:|:|:|
Db 150 LNCRIHKSRNK 161

RESULT 12
US-09-765-111A-16
; Sequence 16, Application US/09765111A
; Patent No. US20020106796A1
; GENERAL INFORMATION:
; APPLICANT: Fletcher, Jonathan A.
; APPLICANT: Kroll, Todd G.
; TITLE OF INVENTION: PAX8-PPARGamma NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: AND POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: B0801/7196/ERP/MAT
; CURRENT APPLICATION NUMBER: US/09/765,111A
; CURRENT FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: US 60/177,109
; PRIOR FILING DATE: 2000-01-20
; PRIOR APPLICATION NUMBER: US 60/225,079
; PRIOR FILING DATE: 2000-08-14
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 16
; LENGTH: 505
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-765-111A-16

Query Match 42.4%; Score 36.5; DB 10; Length 505;
Best Local Similarity 58.3%; Pred. No. 1e+02;
Matches 7; Conservative 4; Mismatches 0; Indels 1; Gaps 1;

QY 3 LNCRLYRKA-NK 13
 |||:|:|:|:|
Db 178 LNCRIHKSRNK 189

RESULT 13
US-10-109-886-6
; Sequence 6, Application US/10109886
; Patent No. US20020119499A1
; GENERAL INFORMATION:
; APPLICANT: TANABE SEIVAKU CO. LTD.
; APPLICANT: MIZUKAMI, Junko
; APPLICANT: TANIGUCHI, Tomoyasu
; TITLE OF INVENTION: METHOD FOR IDENTIFYING OR SCREENING AGONIST AND
; TITLE OF INVENTION: ANTAGONIST TO PPAR
; FILE REFERENCE: TANIGUCHI=6
; CURRENT APPLICATION NUMBER: US/10/109,886
; CURRENT FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 09/514,247
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: PCT/JF98/03734
; PRIOR FILING DATE: 1998-08-24
```



; PRIOR APPLICATION NUMBER: JP231084/1997  
; PRIOR FILING DATE: 1997-08-27  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 6  
; LENGTH: 506  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-109-886-6

Query Match 42.4%; Score 36.5; DB 12; Length 506;  
Best Local Similarity 58.3%; Pred. No. 1e+02;  
Matches 7; Conservative 4; Mismatches 0; Indels 1; Gaps 1;

QY 3 LNCRLYRKA-NK 13  
||||::|: ||  
Db 178 LNCRIHKSRNK 189

## RESULT 14

US-09-765-111A-2  
; Sequence 2, Application US/09765111A  
; Patent No. US20020106796A1  
; GENERAL INFORMATION:  
; APPLICANT: Fletcher, Jonathan A.  
; TITLE OF INVENTION: PAX8-PPARGAMMA NUCLEIC ACID MOLECULES  
; FILE OF INVENTION: AND POLYPEPTIDES AND USES THEREOF  
; FILE REFERENCE: B0801/7196/ERP/MAT  
; CURRENT APPLICATION NUMBER: US/09/765,111A  
; CURRENT FILING DATE: 2001-01-18  
; PRIOR APPLICATION NUMBER: US 60/177,109  
; PRIOR FILING DATE: 2000-01-20  
; PRIOR APPLICATION NUMBER: US 60/225,079  
; PRIOR FILING DATE: 2000-08-14  
; NUMBER OF SEQ ID NOS: 47  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 2  
; LENGTH: 777  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-09-765-111A-2

Query Match 42.4%; Score 36.5; DB 10; Length 777;  
Best Local Similarity 58.3%; Pred. No. 1.6e+02;  
Matches 7; Conservative 4; Mismatches 0; Indels 1; Gaps 1;

QY 3 LNCRLYRKA-NK 13  
||||::|: ||  
Db 450 LNCRIHKSRNK 461

## RESULT 15

US-09-765-111A-23  
; Sequence 23, Application US/09765111A  
; Patent No. US20020106796A1  
; GENERAL INFORMATION:  
; APPLICANT: Fletcher, Jonathan A.  
; TITLE OF INVENTION: PAX8-PPARGAMMA NUCLEIC ACID MOLECULES  
; FILE OF INVENTION: AND POLYPEPTIDES AND USES THEREOF  
; FILE REFERENCE: B0801/7196/ERP/MAT  
; CURRENT APPLICATION NUMBER: US/09/765,111A  
; CURRENT FILING DATE: 2001-01-18  
; PRIOR APPLICATION NUMBER: US 60/177,109  
; PRIOR FILING DATE: 2000-01-20  
; PRIOR APPLICATION NUMBER: US 60/225,079  
; PRIOR FILING DATE: 2000-08-14  
; NUMBER OF SEQ ID NOS: 47  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 23  
; LENGTH: 811  
; TYPE: PRT

; ORGANISM: Homo Sapiens  
US-09-765-111A-23

Query Match 42.4%; Score 36.5; DB 10; Length 811;  
Best Local Similarity 58.3%; Pred. No. 1.7e+02;  
Matches 7; Conservative 4; Mismatches 0; Indels 1; Gaps 1;

QY 3 LNCRLYRKA-NK 13  
||||::|: ||  
Db 484 LNCRIHKSRNK 495

Search completed: December 2, 2002, 10:05:18  
Job time : 2.60697 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 2, 2002, 10:00:35 ; Search time 3.04478 Seconds  
(without alignments)  
536.751 Million cell updates/sec

Title: US-09-741-106-12  
Perfect score: 86  
Sequence: 1 AKLNCLRYKANKSSKL 17  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues  
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_73.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID    | Description        |
|------------|-------|-------------|--------|----------|--------------------|
| 1          | 86    | 100.0       | 464    | 1 XHHU3  | antithrombin III P |
| 2          | 82    | 95.3        | 431    | 1 JX0364 | antithrombin III - |
| 3          | 82    | 95.3        | 433    | 1 A61435 | antithrombin III - |
| 4          | 82    | 95.3        | 465    | 1 S28219 | antithrombin III P |
| 5          | 80    | 93.0        | 465    | 2 I59611 | antithrombin III - |
| 6          | 43    | 50.0        | 2010   | 2 B71616 | phosphatase (acid  |
| 7          | 42    | 48.8        | 407    | 1 A69989 | probable hexosyltr |
| 8          | 41    | 47.7        | 285    | 2 D83380 | hypothetical prote |
| 9          | 41    | 47.7        | 531    | 2 JT0531 | muscarinic acetyl  |
| 10         | 41    | 47.7        | 532    | 2 JT0530 | muscarinic acetyl  |
| 11         | 41    | 47.7        | 664    | 2 T12988 | hypothetical prote |
| 12         | 40    | 46.5        | 244    | 2 D97281 | pseudouridylate sy |
| 13         | 40    | 46.5        | 318    | 1 YXSAT3 | thymidylate synth  |
| 14         | 40    | 46.5        | 623    | 2 E64201 | transport ATP-bind |
| 15         | 40    | 46.5        | 2357   | 2 E59249 | class VII unconven |
| 16         | 39    | 45.3        | 285    | 2 E59905 | transcription regu |
| 17         | 39    | 45.3        | 319    | 2 T48504 | hypothetical prote |
| 18         | 39    | 45.3        | 514    | 2 A49838 | site-specific reco |
| 19         | 39    | 45.3        | 514    | 2 AF1988 | fdx element site-  |
| 20         | 39    | 45.3        | 655    | 2 A12556 | hypothetical prote |
| 21         | 39    | 45.3        | 716    | 2 G44490 | retrovirus-related |
| 22         | 39    | 45.3        | 943    | 2 G81070 | lactoferrin-bindin |
| 23         | 39    | 45.3        | 944    | 2 C81798 | lactoferrin bindin |
| 24         | 39    | 45.3        | 1700   | 2 S08167 | Balbani ring 3 pr  |
| 25         | 39    | 45.3        | 2067   | 2 A42854 | probable spindle p |
| 26         | 38    | 44.2        | 134    | 2 JC6091 | kinetoplast DNA-as |
| 27         | 38    | 44.2        | 173    | 2 A72612 | hypothetical prote |
| 28         | 38    | 44.2        | 257    | 2 E84107 | teichuronic acid b |
| 29         | 38    | 44.2        | 301    | 2 F82287 | transcription regu |

|    |    |      |      |   |        |                    |
|----|----|------|------|---|--------|--------------------|
| 30 | 38 | 44.2 | 310  | 2 | T20208 | hypothetical prote |
| 31 | 38 | 44.2 | 402  | 2 | T29703 | hypothetical prote |
| 32 | 38 | 44.2 | 434  | 2 | E71638 | UDP-glucose 6-dehy |
| 33 | 38 | 44.2 | 473  | 2 | S50755 | hypothetical prote |
| 34 | 38 | 44.2 | 521  | 1 | ACFFNN | nicotinic acetylch |
| 35 | 38 | 44.2 | 547  | 2 | C87992 | protein W09G3.4 [i |
| 36 | 38 | 44.2 | 681  | 2 | S33316 | structural protein |
| 37 | 38 | 44.2 | 681  | 2 | A45705 | type I transmembra |
| 38 | 38 | 44.2 | 725  | 1 | JC5016 | hyaluronan recepto |
| 39 | 38 | 44.2 | 1079 | 2 | T28197 | probable DNA-dirc  |
| 40 | 38 | 44.2 | 1647 | 2 | T41267 | hypothetical prote |
| 41 | 38 | 44.2 | 2718 | 2 | A23475 | G surface protein  |
| 42 | 37 | 43.0 | 134  | 2 | F69056 | ribosomal protein  |
| 43 | 37 | 43.0 | 175  | 2 | D75136 | rubrerythrin PAB05 |
| 44 | 37 | 43.0 | 202  | 2 | A81441 | probable orotate p |
| 45 | 37 | 43.0 | 241  | 2 | B84078 | two-component resp |

ALIGNMENTS

RESULT 1

XHHU3  
antithrombin III precursor [validated] - human  
C:Species: Homo sapiens (man)  
C:Date: 05-Apr-1983 #sequence-revision 05-Apr-1983 #text-change 08-Dec-2000  
C:Accession: A49494; A93453; A92431; A33305; A94445; A34190; S02530; S63600;  
R:Olds, R.J.; Lane, D.A.; Chowdhury, V.; De Stefano, V.; Leone, G.; Thein, S.L.  
Biochemistry 32, 4216-4224, 1993  
A:Title: Complete nucleotide sequence of the antithrombin gene: evidence for homology  
A:Reference number: A49494; MUID:93237227; PMID:8476848  
A:Accession: A49494  
A:Molecule type: DNA  
A:Residues: 1-464 <OLD>  
A:Cross-references: EMBL:X68793; NID:928906; PIDN:CAA48690.1; PID:g28907  
R:Boek, S.C.; Wion, K.L.; Vehar, G.A.; Lawn, R.M.  
Nucleic Acids Res. 10, 8113-8125, 1982  
A:Title: Cloning and expression of the cDNA for human antithrombin III.  
A:Reference number: A93453; MUID:83143280; PMID:6298709  
A:Accession: A93453  
A:Molecule type: mRNA  
A:Residues: 1-464 <BOC1>  
A:Cross-references: GB:L00130; GB:J00102; GB:J00103; GB:J00104; NID:g179138; PIDN:AAB  
R:Chandra, T.; Stackhouse, R.; Kidd, V.J.; Woo, S.L.C.  
Proc. Natl. Acad. Sci. U.S.A. 80, 1845-1848, 1983  
A:Title: Isolation and sequence characterization of a cDNA clone of human antithrombi  
A:Reference number: A93943; MUID:83169777; PMID:6572945  
A:Accession: A93943  
A:Molecule type: mRNA  
A:Residues: 1-464 <CHA>  
A:Cross-references: GB:L00130; GB:J00102; GB:J00103; GB:J00104; NID:g179138; PIDN:AAB  
R:Prochownik, E.V.; Markham, A.F.; Orkin, S.H.  
J. Biol. Chem. 258, 8389-8394, 1983  
A:Title: Isolation of a cDNA clone for human antithrombin III.  
A:Reference number: A92431; MUID:83238456; PMID:6305982  
A:Accession: A92431  
A:Molecule type: mRNA  
A:Residues: 42-96, 'R', 98-464 <PRO>  
A:Note: the authors translated the codon GAC for residue 206 as Asn, GGC for residue  
R:Chang, J.Y.  
J. Biol. Chem. 264, 3111-3115, 1989  
A:Title: Binding of heparin to human antithrombin III activates selective chemical mo  
in III.  
A:Reference number: A33305; MUID:89123426; PMID:2492530  
A:Accession: A33305  
A:Molecule type: protein  
A:Residues: 90-105;124-143;147-161;165-171;268-273;381-402;446-457 <CH2>  
R:Petersen, T.E.; Dudek-Wojciechowska, G.; Sottrup-Jensen, L.; Magnusson, S.  
In The Physiological Inhibitors of Blood Coagulation and Fibrinolysis, Collen, D., W  
A:Title: Primary structure of antithrombin-III (heparin cofactor). Partial homology b  
A:Reference number: A94445  
A:Accession: A94445

A:Molecule type: protein  
A:Residues: 33-68,'QE',71-242:243-245,248-250,254-464 <PET>  
A:Note: carbohydrate-binding site and disulfide bonds  
R:Zetlmeissl, G.; Conradt, H.S.; Nmtz, M.; Karges, H.E.  
J. Biol. Chem. 264, 21153-21159, 1989  
A:Title: Characterization of recombinant human antithrombin III synthesized in Chinese hamster ovary (CHO) cells  
A:Reference number: A34190; MUID:90078215; PMID:2592368  
A:Accession: A34190  
A:Molecule type: protein  
A:Residues: 33-39 <ZET>  
A:Experimental source: recombinant protein from Chinese hamster ovary (CHO) cells  
R:Liu, C.S.; Chang, J.Y.  
Eur. J. Biochem. 167, 247-252, 1987  
A:Title: Probing the heparin-binding domain of human antithrombin III with V8 protease.  
A:Reference number: S02530; MUID:87304255; PMID:3305015  
A:Accession: S02530  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 33-39,'X',41-42:67-77,'XX',80-88,'X',90-92:137-142;189-202;210-214,'X';265-270  
R:Chang, W.S.W.; Wardell, M.R.; Lomas, D.A.; Carrell, R.W.  
Biochem. J. 314, 647-653, 1996  
A:Title: Probing serpin reactive-loop conformations by proteolytic cleavage.  
A:Reference number: S63599; MUID:96239126; PMID:8670081  
A:Accession: S63600  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 414-428 <CH3>  
R:Bock, S.C.; Levitan, D.J.  
Nucleic Acids Res. 11, 8569-8582, 1983  
A:Title: Characterization of an unusual DNA length polymorphism 5' to the human antithrombin III gene.  
A:Reference number: I37191; MUID:84169500; PMID:6672771  
A:Accession: I37191  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-14 <BOC2>  
A:Cross-references: EMBL:X00237; NID:g28917; PIDN:CAA25059.1; PID:g28918; EMBL:X00238; NID:g28919; MUID:84169500; PMID:6672771  
R:Koide, T.; Odani, S.; Takahashi, K.; Ono, T.; Sakuragawa, N.  
Proc. Natl. Acad. Sci. U.S.A. 81, 289-293, 1984  
A:Title: Antithrombin III Toyama: replacement of arginine-47 by cysteine in hereditary antithrombin III.  
A:Reference number: A29371; MUID:84119472; PMID:6582486  
A:Accession: A29371  
A:Molecule type: protein  
A:Residues: 53-68,'QE',71-78,'C',80-135 <KOI>  
A:Note: variant Toyama with a substitution of Cys for Arg-79 and consequently cannot bind heparin.  
R:Bock, S.C.; Marrian, J.A.; Radziejewska, E.  
Biochemistry 27, 6171-6178, 1988  
A:Title: Antithrombin III Utah: proline-407 to leucine mutation in a highly conserved region of the antithrombin III gene.  
A:Reference number: I52399; MUID:89050967; PMID:3191114  
A:Note: a published erratum appears in Biochemistry 28, 3628 1989  
A:Accession: I52399  
A:Status: translation not shown; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 137-208 <BOC3>  
A:Cross-references: GB:M21643; NID:g179147; PIDN:AAA51793.1; PID:g457132  
A:Accession: I65277  
A:Status: translation not shown; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 137-254 <BOC4>  
A:Cross-references: GB:M21644; NID:g179148; PIDN:AAA51794.1; PID:g179151  
A:Accession: I65278  
A:Status: translation not shown; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 407-464 <BOC5>  
A:Cross-references: GB:M21645; NID:g179149; PIDN:AAA51795.1; PID:g179152  
A:Accession: I65279  
A:Status: translation not shown; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-438,'L',440-464 <BOC6>  
A:Cross-references: GB:M21642; NID:g179159; PIDN:AAA51796.1; PID:g179161  
A:Note: mutant Utah  
R:Borg, J.Y.; Brennan, S.O.; Carrell, R.W.; George, P.; Perry, D.J.; Shaw, J.  
FEBS Lett. 266, 163-166, 1990  
A:Title: Antithrombin roven-IV 24 Arg->Cys. The amino-terminal contribution to heparin binding.

A:Reference number: S10716; MUID:90306344; PMID:2365065  
A:Accession: S10716  
A:Molecule type: protein  
A:Residues: 53-61 <BOR>  
A:Note: variant form Rouven-IV, 56-Cys, was also sequenced  
R:Grund, C.B.; Thomas, F.; Millar, D.S.; Krawczak, M.; Mellissari, E.; Lindo, V.; Mof Blood 78, 1027-1032, 1991  
A:Title: Recurrent deletion in the human antithrombin III gene.  
A:Reference number: A44935; MUID:91329813; PMID:1868237  
A:Accession: A44935  
A:Molecule type: DNA  
A:Residues: 271-276,'GRVQHL' <GR1>  
A:Cross-references: GB:S49757; NID:g233568; PIDN:AA19467.1; PID:g233569  
A:Experimental source: antithrombin III deficiency patient 1  
A:Note: sequence extracted from NCBI backbone (NCBIN:49757, NCBIP:49758)  
A:Accession: B44935  
A:Molecule type: DNA  
A:Residues: 271-276,'VVFVSIYDVGQVPLSARG' <GR2>  
A:Cross-references: GB:S49759; NID:g233570; PIDN:AA19468.1; PID:g233571  
A:Experimental source: antithrombin III deficiency patient 2  
A:Note: sequence extracted from NCBI backbone (NCBIN:49759, NCBIP:49760)  
A:Note: different frameshift mutations at an identical site in unrelated patients sug  
R:Daly, M.; Perry, D.J.; Harper, P.L.; Daly, H.M.; Roques, A.W.; Carrell, R.W.  
Thromb. Haemost. 67, 521-525, 1992  
A:Title: Insertions/deletions in the antithrombin gene: 3 mutations associated with n  
A:Reference number: I59610; MUID:92390894; PMID:1325679  
A:Accession: I59610  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 79,'CIGTQVQGPPLCYHLSAPGRFQE' <DAL1>  
A:Cross-references: GB:S43612; NID:g254806; PIDN:AA23132.1; PID:g254807  
A:Note: frameshift mutant, patient 1  
A:Accession: I81229  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 401,'RHFLR' <DAL2>  
A:Cross-references: GB:S43621; NID:g254810; PIDN:AA23134.1; PID:g254811  
A:Note: frameshift mutant, patient 3  
R:Bjork, I.; Danielsson, A.; Fenton II, J.W.; Jornvall, H.  
FEBS Lett. 126, 257-260, 1981  
A:Title: The site in human antithrombin for functional proteolytic cleavage by human  
A:Reference number: A91287; MUID:81212814; PMID:7238875  
A:Contents: annotation; inhibitory site  
R:Blackburn, M.N.; Smith, R.L.; Carson, J.; Sibley, C.C.  
J. Biol. Chem. 259, 939-941, 1984  
A:Title: The heparin-binding site of antithrombin III. Identification of a critical t  
A:Reference number: A92488; MUID:8411578; PMID:6693405  
A:Contents: annotation; heparin-binding site  
C:Genetics:  
A:Gene: GDB:AT3  
A:Cross-references: GDB:119024; OMIM:107300  
A:Map position: 1q23-1q25,1  
A:Introns: 14/2, 136/3; 208/3; 254/3; 385/1; 406/3  
C:Function:  
A:Description: in blood plasma inhibits thrombin and activated coagulation factor X,  
C:Superfamily: antithrombin III  
C:Keywords: acute phase; glycoprotein; heparin binding; plasma; serine proteinase inh  
F:1-32/Domain: signal sequence #status predicted <SIG>  
F:33-464/Product: antithrombin III #status experimental <WAT>  
F:40-160,53-127,279-462/Disulfide bonds: #status experimental  
F:81/Binding site: heparin (Trp) #status experimental  
F:128,167,187,224/Binding site: carbohydrate (Asn) (covalent) #status experimental  
F:425-426/Cleavage site: Arg-Ser (thrombin) #status experimental  
F:425/Inhibitory site: Arg (thrombin, coagulation factor Xa) #status experimental

Query Match 100.0%; Score 86; DB 1; Length 464;

Best Local Similarity 100.0%; Pred. No. 5.8e-07; Mismatches 0; Indels 0; Gaps 0;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKLNCRLYRKANKSKSL 17

|||||

Db 156 AKLNCRLYRKANKSKSL 172

## RESULT 2

JX0364  
antithrombin III - pig  
C:Species: Sus scrofa domestica (domestic pig)  
C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C:Accession: JX0364  
R:Tokunaga, F.; Goto, T.; Wakabayashi, S.; Koide, T.  
J. Biochem. 116, 1164-1170, 1994  
A:Title: Amino acid sequence of porcine antithrombin III.  
A:Reference number: JX0364; MUID:95204393; PMID:7896748  
A:Accession: JX0364  
A:Molecule type: protein  
A:Residues: 1-431 <TOK>  
C:Function:  
C:Description: inhibits in blood plasma thrombin and activated coagulation factor X, etc  
C:Superfamily: antithrombin III  
C:Keywords: glycoprotein; heparin binding  
F:378-382/Region: hinge #status predicted  
F:9-129,22-96,248-431/Disulfide bonds: #status predicted  
F:136,156,193/Binding site: carbohydrate (Asn) (covalent) #status experimental  
F:394/Inhibitory site: Arg (thrombin) #status predicted

Query Match 95.3%; Score 82; DB 1; Length 431;  
Best Local Similarity 94.1%; Pred. No. 2.7e-06;  
Matches 16; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKLNCRLYRKANKSSKL 17  
|||||  
DB 125 AKLNCRLYRKANKSSEL 141

## RESULT 3

A61435  
antithrombin III - bovine  
C:Species: Bos primigenius taurus (cattle)  
C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C:Accession: A61435  
R:Mejdoub, H.; Le Ret, M.; Boulanger, Y.; Maman, M.; Choay, J.; Reinbolt, J.  
J. Protein Chem. 10, 205-212, 1991  
A:Title: The complete amino acid sequence of bovine antithrombin (ATIII).  
A:Reference number: A61435; MUID:92029517; PMID:1930634  
A:Accession: A61435  
A:Molecule type: protein  
A:Residues: 1-433 <MEJ>  
C:Comment: This serpin inhibits thrombin and other clotting factors in the presence of H  
C:Function:  
C:Description: inhibits in blood plasma thrombin and activated coagulation factor X, etc  
C:Superfamily: antithrombin III  
C:Keywords: anticoagulant; glycoprotein; heparin binding; plasma; serine proteinase inh  
F:9-129,22-96,248-431/Disulfide bonds: #status predicted  
F:97,136,156,193/Binding site: carbohydrate (Asn) (covalent) #status experimental  
F:394/Inhibitory site: Arg (thrombin) #status experimental

Query Match 95.3%; Score 82; DB 1; Length 433;  
Best Local Similarity 94.1%; Pred. No. 2.7e-06;  
Matches 16; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKLNCRLYRKANKSSKL 17  
|||||  
DB 125 AKLNCRLYRKANKSSEL 141

## RESULT 4

S28219  
antithrombin III precursor - sheep  
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)  
C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C:Accession: S28219  
R:Nielsen, R.W.L.M.; Sturk, A.; Hordijk, P.L.; Michiels, F.; Peters, M.  
Biochim. Biophys. Acta 1171, 207-210, 1992  
A:Title: Sequence characterization of a sheep cDNA for antithrombin III.  
A:Reference number: S28219; MUID:93129622; PMID:1482684

A:Accession: S28219  
A:Molecule type: mRNA  
A:Residues: 1-465 <NIE>  
A:Cross-references: EMBL:X68287; NID:g1194; PIDN:CAA48347.1; PID:g1195  
C:Function:  
C:Description: inhibits in blood plasma thrombin and activated coagulation factor X,  
C:Superfamily: antithrombin III  
C:Keywords: glycoprotein; plasma; serine proteinase inhibitor  
F:1-32/Domain: signal sequence #status predicted <SIG>  
F:33-465/Product: antithrombin III #status predicted <MAT>  
F:129,168,188,225/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 95.3%; Score 82; DB 1; Length 465;  
Best Local Similarity 94.1%; Pred. No. 2.8e-06;  
Matches 16; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKLNCRLYRKANKSSKL 17  
|||||  
DB 157 AKLNCRLYRKANKSSEL 173

## RESULT 5

I59611  
antithrombin III - mouse  
C:Species: Mus sp. (mouse)  
C>Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 16-Jul-1999  
C:Accession: I59611  
R:Wu, J.K.; Sheffield, W.P.; Blajchman, M.A.  
Thromb. Haemost. 68, 291-296, 1992  
A:Title: Molecule cloning and cell-free expression of mouse antithrombin III.  
A:Reference number: I59611; MUID:93069082; PMID:1440494  
A:Accession: I59611  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-465 <RES>  
A:Cross-references: GB:S47225; NID:g258958; PIDN:AAB23965.1; PID:g258959  
C:Superfamily: antithrombin III

Query Match 93.0%; Score 80; DB 2; Length 465;  
Best Local Similarity 94.1%; Pred. No. 6.3e-06;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AKLNCRLYRKANKSSKL 17  
|||||  
DB 157 AKLNCRLYRKANKSSDL 173

## RESULT 6

B71616  
phosphatase (acid phosphatase family) PFB0380c - malaria parasite (Plasmodium falcipa  
C:Species: Plasmodium falciparum  
C>Date: 13-Nov-1998 #sequence\_revision 13-Nov-1998 #text\_change 21-Jul-2000  
C:Accession: B71616  
R:Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.  
; Pertea, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H  
Science 282, 1126-1132, 1998  
A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.  
A:Reference number: A71600; MUID:99021743; PMID:9804551  
A:Accession: B71616  
A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-2010 <GAR>  
A:Cross-references: GB:AE001391; GB:AE001362; NID:g3845168; PIDN:AAC71865.1; PID:g384

## A:Gene: PFB0380c

Query Match 50.0%; Score 43; DB 2; Length 2010;  
Best Local Similarity 57.1%; Pred. No. 52;  
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 LNCRLYRKANKSSK 16  
:| ||| | | | :|

**RESULT 7**

A69989  
probable hexosyltransferase (EC 2.4.1.-) ytcC - Bacillus subtilis  
N:Alternate names: lipopolysaccharide N-acetylglucosaminyltransferase ytcC  
C:Species: Bacillus subtilis  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 16-Jun-2000  
C:Accession: A69989  
Nature 390, 249-256, 1997  
A:A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Gallizzi, A.; Galler,  
iech, J.; Harwood, C.R.; Henaute, A.; Hilbert, H.; Holtsappel, S.; Hosono, S.; Hulio, M.F.;  
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,  
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mausel,  
Y., M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohll, T.M.; Portetelle,  
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon,  
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seroh,  
akeuchi, M.; Tanakashi, A.; Tanaka, T.; Terpstora, P.; Togonoi, A.; Tosato, V.; Uchiyama,  
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.  
A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.

A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.  
A:Reference number: A69580; MUID:98044033; PMID:9384377  
A:Accession: A69580  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-407 <GUN>  
A:Cross-references: GB:Z99119; GB:AL009126; NID:g2635411; PIDN:CAB15066.1; PID:g2635572  
A:Experimental source: strain 168  
C:Genetics:  
A:Gene: ytcC  
C:Superfamily: probable hexosyltransferase ytcX  
C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 48.8%; Score 42; DB 1; Length 407;  
Best Local Similarity 53.3%; Pred. No. 19;  
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

**QY 3 LNCRLYRKANKSKL 17**  
|||:::||||  
**Db 90 LNVSLYKAKSPNSKI 104**

**RESULT 8**  
hypothetical protein PA2133 [imported] - Pseudomonas aeruginosa (strain PAO1)  
C:Species: Pseudomonas aeruginosa  
C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
C:Accession: D83380  
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Bd,  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,  
.; Lory, S.; Olson, M.V.  
Nature 406, 955-964, 2000  
A:A:Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic patho  
A:Reference number: AB2950; MUID:20437337; PMID:10984043  
A:Accession: D83380  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-285 <STO>  
A:Cross-references: GB:AE004640; GB:AE004091; NID:g9948139; PIDN:AAG05521.1; GSPDB:GN001  
A:Experimental source: strain PAO1  
C:Genetics:  
A:Gene: PA2133

Query Match 47.7%; Score 41; DB 2; Length 285;  
Best Local Similarity 46.7%; Pred. No. 21;  
Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

**QY 2 KLNCRLYRKANKSSK 16**  
||:|||||::|:

F:105-126/Domain: transmembrane #status predicted <TM3>  
 F:147-169/Domain: transmembrane #status predicted <TM4>  
 F:192-214/Domain: transmembrane #status predicted <TM5>  
 F:444-464/Domain: transmembrane #status predicted <TM6>  
 F:479-498/Domain: transmembrane #status predicted <TM7>  
 F:8,13/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 47.7%; Score 41; DB 2; Length 532;  
 Best Local Similarity 50.0%; Pred. No. 36;  
 Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 3 LNCRLYRKANKSK 16  
 | ||| || | | | |  
 Db 212 LYCRIVRETEKRTK 225

## RESULT 11

T12988

hypothetical protein T21L8.80 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cross)

C>Date: 13-Aug-1999 #sequence\_revision 13-Aug-1999 #text\_change 20-Jun-2000

C:Accession: T12988

R:Choisne, N.; Robert, C.; Brottier, P.; Wincker, P.; Cattolico, L.; Artiguenave, F.; Sa

submitted to the Protein Sequence Database, July 1999

A:Reference number: Z17586

A:Accession: T12988

A:Molecule type: DNA

A:Residues: 1-664 <CHO>

A:Cross-references: EMBL:AL096860; GSPDB:GN00061; ATP:T21L8.80

A:Experimental source: cultivar Columbia; BAC clone T21L8

C:Genetics:

A:Gene: ATP:T21L8.80

A:Map position: 3

A:Introns: 106/3; 374/3; 427/3; 606/1

C:Superfamily: Arabidopsis thaliana hypothetical protein F18N11.140

Query Match 47.7%; Score 41; DB 2; Length 664;

Best Local Similarity 66.7%; Pred. No. 44;

Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 LNCRLYRKANKS 14  
 ||||| | | | |  
 Db 245 LNCREYPKNNKT 256

## RESULT 12

D97281

pseudouridylate synthase, TRUA [imported] - Clostridium acetobutylicum

C:Species: Clostridium acetobutylicum

C>Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 30-Sep-2001

C:Accession: D97281

R:Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,

; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.

J. Bacteriol. 183, 4823-4838, 2001

A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cl

A:Reference number: A96900; MUID:21359325; PMID:21359325

A:Accession: D97281

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-244 <KUR>

A:Cross-references: GB:AE001437; PIDN:AAK81039.1; PID:gl5026164; GSPDB:GN00168

A:Experimental source: Clostridium acetobutylicum ATCC824

C:Genetics:

A:Gene: CAC3099

C:Superfamily: tRNA-pseudouridine synthase I

Query Match 46.5%; Score 40; DB 2; Length 244;

Best Local Similarity 66.7%; Pred. No. 27;

Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 KLCRLYRKANK 13  
 ||||| | | | |  
 Db 134 KLCMDLMREATK 145

## RESULT 13

YASAT3

thymidylate synthase (EC 2.1.1.45) - Staphylococcus aureus plasmid pSK1 transposon Tn

C:Species: Staphylococcus aureus

C>Date: 31-Mar-1993 #sequence\_revision 31-Mar-1993 #text\_change 11-Jun-1999

C:Accession: S04163; S14178

R:Rouch, D.A.; Messerotti, L.J.; Loo, L.S.L.; Jackson, C.A.; Skurray, R.A.

Mol. Microbiol. 3, 161-175, 1989

A:Title: Trimethoprim resistance transposon Tn4003 from Staphylococcus aureus encodes

A:Reference number: S04162; MUID:89343620; PMID:2548057

A:Accession: S04163

A:Molecule type: DNA

A:Residues: 1-318 <ROU>

A:Cross-references: EMBL:X13290; NID:g46747; PIDN:CAA31648.1; PID:g46749

R:Burdaska, A.; Ott, M.; Bannwarth, W.; Then, R.L.

FEBS Lett. 266, 159-162, 1990

A:Title: Identical genes for trimethoprim-resistant dihydrofolate reductase from Stap

A:Reference number: S10715; MUID:90306343; PMID:2365064

A:Accession: S14178

A>Status: preliminary; translation not shown

A:Molecule type: DNA

A:Residues: 253-318 <BUR>

A:Cross-references: EMBL:Y07536; NID:g46551; PIDN:CAA68823.1; PID:g46552

C:Genetics:

A:Gene: thvA

A:Genome: plasmid

C:Superfamily: thymidylate synthase; thymidylate synthase homology

C:Keywords: deoxyribonucleotide biosynthesis; methyltransferase

F:6-318/Domain: thymidylate synthase homology <TDS>

F:201/Active site: Cys #status predicted

Query Match 46.5%; Score 40; DB 1; Length 318;

Best Local Similarity 60.0%; Pred. No. 34;

Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 KLCRLYRKA 11  
 ||||| | | | |  
 Db 213 KLCQLYORS 222

## RESULT 14

E64201

transport ATP-binding protein msbA homolog MG014 - Mycoplasma genitalium

C:Species: Mycoplasma genitalium

C>Date: 17-Nov-1995 #sequence\_revision 17-Nov-1995 #text\_change 02-Feb-2001

C:Accession: E64201

R:Fraser, C.M.; Gocayne, J.D.; White, O.; Adams, M.D.; Clayton, R.A.; Fleischmann, R.

M.; Fuhrmann, J.; Nguyen, D.; Utterback, T.R.; Saudek, D.M.; Phillips, C.A.; Merrick,

C.A.; Venter, J.C.

Science 270, 397-403, 1995

A:Title: The minimal gene complement of Mycoplasma genitalium.

A:Reference number: A64200; MUID:96026346; PMID:7569993

A:Accession: E64201

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-623 <TIGR>

A:Cross-references: GB:U39680; GB:L43967; NID:g3844621; PIDN:AA71230.1; PID:gl045683

A:Experimental source: strain G-37

C:Genetics:

A:Genetic code: SGC3

C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homolog

C:Keywords: ATP; nucleotide binding; P-loop

F:383-587/Domain: ATP-binding cassette homology <ABC>

F:400-407/Region: nucleotide-binding motif A (P-loop)

Query Match 46.5%; Score 40; DB 2; Length 623;

Best Local Similarity 53.8%; Pred. No. 62;

Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 4 NCRLYRKANKSK 16  
 || ||| | | | |





GenCore version 5.1.1.3

Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 2, 2002, 09:59:55 ; Search time 1.69154 Seconds  
(without alignments)  
416.837 Million cell updates/sec

Title: US-09-741-106-12

Perfect score: 86

Sequence: 1 AKLNCRLYRKANKSKL 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB ID        | Description          |
|------------|-------|-------------|--------|--------------|----------------------|
| 1          | 86    | 100.0       | 464    | 1 ANT3_HUMAN | P01008 homo sapien   |
| 2          | 82    | 95.3        | 433    | 1 ANT3_BOVIN | P41361 bos taurus    |
| 3          | 82    | 95.3        | 465    | 1 ANT3_SHEEP | P32262 ovis aries    |
| 4          | 80    | 93.0        | 465    | 1 ANT3_MOUSE | P32261 mus musculus  |
| 5          | 41    | 47.7        | 531    | 1 ACW5_RAT   | P08911 rattus norv   |
| 6          | 41    | 47.7        | 532    | 1 ACW5_HUMAN | P08912 homo sapien   |
| 7          | 41    | 47.7        | 532    | 1 ACW5_MACMU | P56490 macaca mula   |
| 8          | 40    | 46.5        | 318    | 1 TVSY_STRAU | P13954 staphylococ   |
| 9          | 40    | 46.5        | 623    | 1 Y014_MYCGE | P47260 mycoplasma    |
| 10         | 39    | 45.3        | 285    | 1 Y0FA_BACSU | O34685 bacillus su   |
| 11         | 39    | 45.3        | 869    | 1 P021_SCICO | Q03279 sciatara copr |
| 12         | 39    | 45.3        | 943    | 1 LBPA_NEIMB | Q06379 neisseria m   |
| 13         | 39    | 45.3        | 944    | 1 LBPA_NEIMB | O91tk4 neisseria m   |
| 14         | 39    | 45.3        | 1700   | 1 BAR3_CHITE | Q03376 chironomus    |
| 15         | 39    | 45.3        | 2067   | 1 BTMB_EMENI | P33144 emeritella    |
| 16         | 38    | 44.2        | 375    | 1 GDF8_SHEEP | O18830 ovis aries    |
| 17         | 38    | 44.2        | 434    | 1 UDG_RICPR  | O05973 rickettsia    |
| 18         | 38    | 44.2        | 521    | 1 ACH3_DROME | P04755 drosophila    |
| 19         | 38    | 44.2        | 681    | 1 VGP_MABVM  | P35253 marburg vir   |
| 20         | 38    | 44.2        | 681    | 1 VGP_MABVP  | P35254 marburg vir   |
| 21         | 38    | 44.2        | 724    | 1 HMNR_HUMAN | O75330 homo sapien   |
| 22         | 38    | 44.2        | 2715   | 1 G156_PARPR | P13837 parametech    |
| 23         | 37    | 43.0        | 134    | 1 RS15_METHH | O27474 methanobact   |
| 24         | 37    | 43.0        | 202    | 1 PYRE_CAMJE | P49018 saccharomyc   |
| 25         | 37    | 43.0        | 411    | 1 GP18_YEAST | P46236 fusarium ox   |
| 26         | 37    | 43.0        | 462    | 1 GUNB_FUSOX | P24045 gallus gall   |
| 27         | 37    | 43.0        | 488    | 1 GAB4_CHICK | O14186 schizosacch   |
| 28         | 37    | 43.0        | 846    | 1 YDSB_SCHPO | P87137 schizosacch   |
| 29         | 37    | 43.0        | 929    | 1 YDM6_SCHPO | Q12704 schizosacch   |
| 30         | 37    | 43.0        | 1018   | 1 DPOG_SCHPO | O8wm16 macaca mula   |
| 31         | 36.5  | 42.4        | 104    | 1 NXPI_MACMU | O61200 mus musculus  |
| 32         | 36.5  | 42.4        | 253    | 1 NXPI_MOUSE | P58417 homo sapien   |
| 33         | 36.5  | 42.4        | 271    | 1 NXPI_HUMAN |                      |

|    |      |      |      |   |            |                     |
|----|------|------|------|---|------------|---------------------|
| 34 | 36.5 | 42.4 | 271  | 1 | NXP1_RAT   | Q63366 rattus norv  |
| 35 | 36.5 | 42.4 | 475  | 1 | PPAT_CRIGR | P57797 cricetus     |
| 36 | 36.5 | 42.4 | 475  | 1 | PPAT_RABIT | O19052 oryctolagus  |
| 37 | 36.5 | 42.4 | 477  | 1 | PPAT_XENLA | P37234 xenopus lae  |
| 38 | 36.5 | 42.4 | 504  | 1 | PPAT_PIG   | O62807 sus scrofa   |
| 39 | 36.5 | 42.4 | 505  | 1 | PPAT_BOVIN | O18971 bos taurus   |
| 40 | 36.5 | 42.4 | 505  | 1 | PPAT_HUMAN | P37231 homo sapien  |
| 41 | 36.5 | 42.4 | 505  | 1 | PPAT_MACMU | O18924 macaca mula  |
| 42 | 36.5 | 42.4 | 505  | 1 | PPAT_MOUSE | P37238 mus musculus |
| 43 | 36.5 | 42.4 | 505  | 1 | PPAT_RAT   | O68275 rattus norv  |
| 44 | 36.5 | 42.4 | 1002 | 1 | VGNM_CPSMV | P31630 cowpea seve  |
| 45 | 36   | 41.9 | 212  | 1 | ECML_YEAST | P39715 saccharomyc  |

## ALIGNMENTS

|            |                                                                                                                             |           |      |         |
|------------|-----------------------------------------------------------------------------------------------------------------------------|-----------|------|---------|
| RESULT 1   |                                                                                                                             |           |      |         |
| ANT3_HUMAN |                                                                                                                             |           |      |         |
| ID         | ANT3_HUMAN                                                                                                                  | STANDARD; | PRT; | 464 AA. |
| AC         | P01008;                                                                                                                     |           |      |         |
| DT         | 21-JUL-1986 (Rel. 01, Created)                                                                                              |           |      |         |
| DT         | 21-JUL-1986 (Rel. 01, Last sequence update)                                                                                 |           |      |         |
| DT         | 15-JUN-2002 (Rel. 41, Last annotation update)                                                                               |           |      |         |
| DE         | Antithrombin-III precursor (ATIII) (PRO0309).                                                                               |           |      |         |
| GN         | SERPINC1 OR AT3.                                                                                                            |           |      |         |
| OS         | Homo sapiens (Human).                                                                                                       |           |      |         |
| OC         | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;                                                           |           |      |         |
| OC         | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.                                                                  |           |      |         |
| OX         | NCBI_TaxID=9606;                                                                                                            |           |      |         |
| [1]        |                                                                                                                             |           |      |         |
| RP         | SEQUENCE FROM N.A.                                                                                                          |           |      |         |
| RX         | MEDLINE=83143280; PubMed=6298709;                                                                                           |           |      |         |
| RA         | Bock S.C., Wion K.L., Vohar G.A., Lawn R.M.;                                                                                |           |      |         |
| RT         | "Cloning and expression of the cDNA for human antithrombin III.;"                                                           |           |      |         |
| RL         | Nucleic Acids Res. 10:8113-8125(1982).                                                                                      |           |      |         |
| [2]        |                                                                                                                             |           |      |         |
| RP         | SEQUENCE FROM N.A.                                                                                                          |           |      |         |
| RX         | MEDLINE=83169777; PubMed=6572945;                                                                                           |           |      |         |
| RA         | Chandra T., Stackhouse R., Kidd V.J., Woo S.L.C.;                                                                           |           |      |         |
| RT         | "Isolation and sequence characterization of a cDNA clone of human antithrombin III.;"                                       |           |      |         |
| RL         | Proc. Natl. Acad. Sci. U.S.A. 80:1845-1848(1983).                                                                           |           |      |         |
| [3]        |                                                                                                                             |           |      |         |
| RP         | SEQUENCE FROM N.A.                                                                                                          |           |      |         |
| RX         | MEDLINE=93237227; PubMed=8476848;                                                                                           |           |      |         |
| RA         | Olds R.J., Lane D.A., Chowdhury V., de Stefano V., Leone G.,                                                                |           |      |         |
| RA         | Thein S.L.;                                                                                                                 |           |      |         |
| RT         | "Complete nucleotide sequence of the antithrombin gene: evidence for homologous recombination causing thrombophilia.;"      |           |      |         |
| RL         | Biochemistry 32:4216-4224(1993).                                                                                            |           |      |         |
| [4]        |                                                                                                                             |           |      |         |
| RP         | SEQUENCE FROM N.A.                                                                                                          |           |      |         |
| RC         | TISSUE=Fetal liver;                                                                                                         |           |      |         |
| RA         | Zhang C., Yu Y., Zhang S., Wei H., Bi J., Zhou G., Dong C., Zai Y.,                                                         |           |      |         |
| RA         | Xu W., Gao F., Liu M., He F.;                                                                                               |           |      |         |
| RT         | "Functional prediction of the coding sequences of 75 new genes deduced by analysis of cDNA clones from human fetal liver.;" |           |      |         |
| RL         | Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.                                                                     |           |      |         |
| [5]        |                                                                                                                             |           |      |         |
| RP         | SEQUENCE FROM N.A., AND VARIANTS GLU-30 AND ALA-147.                                                                        |           |      |         |
| RA         | Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L., Yi Q.,                                                      |           |      |         |
| RA         | Nickerson D.A.;                                                                                                             |           |      |         |
| RL         | Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.                                                                     |           |      |         |
| [6]        |                                                                                                                             |           |      |         |
| RP         | SEQUENCE OF 42-464 FROM N.A.                                                                                                |           |      |         |
| RX         | MEDLINE=83238456; PubMed=6305982;                                                                                           |           |      |         |
| RA         | Prochownik E.V., Markham A.F., Orkin S.H.;                                                                                  |           |      |         |
| RT         | "Isolation of a cDNA clone for human antithrombin III.;"                                                                    |           |      |         |
| RL         | J. Biol. Chem. 258:8389-8394(1983).                                                                                         |           |      |         |
| [7]        |                                                                                                                             |           |      |         |
| RP         | SEQUENCE OF 33-464, CARBOHYDRATE-LINKAGE SITES, AND DISULFIDE BONDS.                                                        |           |      |         |
| RA         | Petersen T.E., Dudek-wojciechowska G., Sottrup-Jensen L.,                                                                   |           |      |         |

- RA Magnusson S.;  
 RT "Primary structure of antithrombin-III (heparin cofactor). Partial  
 RT homology between alpha-1-antitrypsin and antithrombin-III.";  
 RL (in) Collen D., Wiman B., Verstraete M. (eds.);  
 RL The physiological inhibitors of blood coagulation and fibrinolysis,  
 RL pp.43-54, Elsevier, Amsterdam (1979).  
 RN [8]  
 RN ACTIVE SITE.  
 RP MEDLINE=81212814; PubMed=7238875;  
 RX Bjoerk I., Danielsson A., Fenton J.W. II, Joernvall H.;  
 RA "The site in human antithrombin for functional proteolytic cleavage  
 RT by human thrombin.";  
 RL FEBS Lett. 126:257-260(1981).  
 RN [9]  
 RN HEPARIN-BINDING SITE.  
 RP MEDLINE=84111578; PubMed=6693405;  
 RX Blackburn M.N., Smith R.L., Carson J., Sibley C.C.;  
 RA "The heparin-binding site of antithrombin III. Identification of a  
 RT critical tryptophan in the amino acid sequence.";  
 RL J. Biol. Chem. 259:939-941(1984).  
 RN [10]  
 RN MUTAGENESIS OF ALA-414.  
 RP MEDLINE=91192143; PubMed=2013320;  
 RX Austin R.C., Rachubinski R.A., Blajchman M.A.;  
 RA "Site-directed mutagenesis of alanine-382 of human antithrombin III.";  
 RL FEBS Lett. 280:254-258(1991).  
 RN [11]  
 RN X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS).  
 RP MEDLINE=94373498; PubMed=8087553;  
 RX Carrell R.W., Stein P.E., Fermi G., Wardell M.R.;  
 RA "Biological implications of a 3 A structure of dimeric antithrombin.";  
 RL Structure 2:257-270(1994).  
 RN [12]  
 RN X-RAY CRYSTALLOGRAPHY (3.2 ANGSTROMS).  
 RP MEDLINE=95384753; PubMed=7656006;  
 RX Schroeder H.A., de Boer B., DiJkema R., Mulders J., Theunissen H.J.M.,  
 RA Grootenhuys P.D.J., Hol W.G.J.;  
 RL "The intact and cleaved human antithrombin III complex as a model for  
 RT serpin-proteinase interactions.";  
 RL Nat. Struct. Biol. 1:48-54(1994).  
 RN [13]  
 RN X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).  
 RP MEDLINE=97220238; PubMed=9067613;  
 RX Skinner R., Abrahams J.-P., Whistock J.C., Lesk A.M., Carrel R.W.,  
 RA Wardell M.R.;  
 RL "The 2.6 A structure of antithrombin indicates a conformational  
 RT change at the heparin binding site.";  
 RL J. Mol. Biol. 266:601-609(1997).  
 RN [14]  
 RN X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).  
 RP MEDLINE=98437366; PubMed=9761669;  
 RX Skinner R., Chang W.S.W., Jin L., Pei X.Y., Huntington J.A.,  
 RA Abrahams J.-P., Carrell R.W., Lomas D.A.;  
 RL "Implications for function and therapy of a 2.9 A structure of  
 RT binary-complexed antithrombin.";  
 RL J. Mol. Biol. 283:9-14(1998).  
 RN [15]  
 RN REVIEW.  
 RP MEDLINE=91129302; PubMed=2126464;  
 RX Mourey L., Samama J.-P., Delarue M., Choay J., Lormeau J.C.,  
 RA Petitou M., Moras D.;  
 RL "Antithrombin III: structural and functional aspects.";  
 RL Biochimie 72:599-608(1990).  
 RN [16]  
 RN REVIEW ON VARIANTS.  
 RP MEDLINE=94054329; PubMed=8236149;  
 RX Lane D.A., Olds R.J., Boisclair M., Chowdhury V., Thein S.L.,  
 RA Cooper D.N., Blajchman M., Perry D., Emmerich J., Alach M.;  
 RL "Antithrombin III mutation database: first update. For the Thrombin  
 RT and its Inhibitors Subcommittee of the Scientific and Standardization  
 RT Committee of the International Society on Thrombosis and  
 RL Haemostasis.";  
 RL Thromb. Haemost. 70:361-369(1993).
- RA REVIEW ON VARIANTS.  
 RP MEDLINE=95269065; PubMed=7749926;  
 RX Stein P.E., Carrell R.W.;  
 RA "What do dysfunctional serpins tell us about molecular mobility and  
 RT disease?";  
 RL Nat. Struct. Biol. 2:96-113(1995).  
 RN [18]  
 RN REVIEW ON VARIANTS.  
 RP MEDLINE=96263733; PubMed=8664906;  
 RX Perry D.J., Carrell R.W.;  
 RA "Molecular genetics of human antithrombin deficiency.";  
 RL Hum. Mutat. 7:7-22(1996).  
 RN [19]  
 RN VARIANT UTAH.  
 RP MEDLINE=89050967; PubMed=3191114;  
 RX Bock S.C., Marrinan J.A., Radziejewska E.;  
 RA "Antithrombin III Utah: proline-407 to leucine mutation in a highly  
 RT conserved region near the inhibitor reactive site.";  
 RL Biochemistry 27:6171-6178(1988).  
 RN [20]  
 RN VARIANT TOYAMA.  
 RP MEDLINE=84119472; PubMed=6582486;  
 RX Koide T., Odani S., Takahashi K., Ono T., Sakuragawa N.;  
 RA "Antithrombin III Toyama: replacement of arginine-47 by cysteine in  
 RT hereditary abnormal antithrombin III that lacks heparin-binding  
 RL ability.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 81:289-293(1984).  
 RN [21]  
 RN VARIANT BASEL.  
 RP MEDLINE=86111754; PubMed=3080419;  
 RX Chang J.Y., Tran T.H.;  
 RA "Antithrombin III Basel. Identification of a Pro-Leu substitution in  
 RT a hereditary abnormal antithrombin with impaired heparin cofactor  
 RL activity.";  
 RL J. Biol. Chem. 261:1174-1176(1986).  
 RN [22]  
 RN VARIANT DENVER.  
 RP MEDLINE=87109210; PubMed=3805013;  
 RX Stephens A.W., Thalley B.S., Hirs C.H.W.;  
 RA "Antithrombin-III Denver, a reactive site variant.";  
 RL J. Biol. Chem. 262:1044-1048(1987).  
 RN [23]  
 RN VARIANT HAMILTON.  
 RP MEDLINE=89027076; PubMed=3179438;  
 RX Devrak-Kizuk R., Chui D.H.K., Prochownik E.V., Carter C.J.,  
 RA Ofosu F.A., Blajchman M.A.;  
 RL "Antithrombin-III-Hamilton: a gene with a point mutation (guanine to  
 RT adenine) in codon 382 causing impaired serine protease reactivity.";  
 RL Blood 72:1518-1523(1988).  
 RN [24]  
 RN VARIANTS GLASGOW AND NORTHWICK-PARK.  
 RP MEDLINE=88186869; PubMed=3162733;  
 RX Erdjument H., Lamed D.A., Panico M., di Marzo V., Morris H.R.;  
 RA "Single amino acid substitutions in the reactive site of antithrombin  
 RT leading to thrombosis. Congenital substitution of arginine 393 to  
 RL cysteine in antithrombin Northwick Park and to histidine in  
 RL antithrombin Glasgow.";  
 RL J. Biol. Chem. 263:5589-5593(1988).  
 RN [25]  
 RN VARIANT CHICAGO.  
 RP MEDLINE=89388698; PubMed=2781509;  
 RX Erdjument H., Lane D.A., Panico M., di Marzo V., Morris H.R.,  
 RA Bauer K., Rosenberg R.D.;  
 RL "Antithrombin Chicago, amino acid substitution of arginine 393 to  
 RT histidine.";  
 RL Thromb. Res. 54:613-619(1989).  
 RN [26]  
 RN VARIANT ROUEN-4.  
 RP MEDLINE=90306344; PubMed=2365065;  
 RX Borg J.Y., Brennan S.O., Carrell R.W., George P., Perry D.J., Shaw J.;  
 RA "Antithrombin Rouen-IV 24 Arg-->Cys. The amino-terminal contribution  
 RT to heparin binding.";

Query Match 100.0%; Score 86; DB 1; Length 464;  
 Best Local Similarity 100.0%; Pred. No. 2e-07;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKLNCRLYRKANKSSKL 17  
 |||||||||||||||  
 Db 156 AKLNCRLYRKANKSSKL 172

## RESULT 2

ANT3\_BOVIN STANDARD; PRT; 433 AA.  
 AC P41361;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Antithrombin-III (AIII).  
 GN SERPINC1 OR AT3.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euthera; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=92029517; PubMed=1930634;  
 RA Mejdoub H., le Ret M., Boulanger Y., Maman M., Choay J., Reinbolt J.;  
 RT "The complete amino acid sequence of bovine antithrombin (AIII).";  
 RL J. Protein Chem. 10:205-212(1991).  
 RN [2]  
 RP X-RAY CRYSTALLOGRAPHY (3.2 ANGSTROMS).  
 RX MEDLINE=93323115; PubMed=8331659;  
 RA Mourey L., Samama J.-P., Delarue M., Petitou M., Choay J., Moras D.;  
 RT "Crystal structure of cleaved bovine antithrombin III at 3.2-A  
 resolution.";  
 RL J. Mol. Biol. 232:223-241(1993).

CC -!- FUNCTION: MOST IMPORTANT SERINE PROTEASE INHIBITOR IN PLASMA  
 CC THAT REGULATES THE BLOOD COAGULATION CASCADE. AT-III INHIBITS  
 CC THROMBIN AS WELL AS FACTORS IXA, XA AND XIA. ITS INHIBITORY  
 CC ACTIVITY IS GREATLY ENHANCED IN THE PRESENCE OF HEPARIN.  
 CC -!- SUBCELLULAR LOCATION: Extracellular.  
 CC -!- TISSUE SPECIFICITY: PLASMA.  
 CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY.  
 DR PIR; A61435; A61435.  
 DR PDB; IAT; 31-JUL-94.  
 DR InterPro; IPR000215; Serpin.  
 DR Pfam; PF00079; serpin; 1.  
 DR SMART; SM00093; SERPIN; 1.  
 DR PROSITE; PS00284; SERPIN; 1.  
 KW Serpin; Serine protease inhibitor; Glycoprotein; Heparin-binding;  
 KW Plasma; Blood coagulation; 3D-structure.  
 FT BINDING 50 50 HEPARIN (BY SIMILARITY).  
 FT BINDING 130 130 HEPARIN (BY SIMILARITY).  
 FT BINDING 146 146 HEPARIN (BY SIMILARITY).  
 FT ACT\_SITE 394 395 HEPARIN (BY SIMILARITY).  
 FT DISULFID 9 129 REACTIVE BOND.  
 FT DISULFID 22 96  
 FT DISULFID 248 431 N-LINKED (GLCNAC. . .)  
 FT CARBOHYD 97 97 N-LINKED (GLCNAC. . .)  
 FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .)  
 FT CARBOHYD 156 156 N-LINKED (GLCNAC. . .)  
 FT CARBOHYD 193 193 N-LINKED (GLCNAC. . .)  
 SQ SEQUENCE 433 AA; 49126 MW; F6295B3F0195E9A9 CRC64;

Query Match 95.3%; Score 82; DB 1; Length 433;  
 Best Local Similarity 94.1%; Pred. No. 9.3e-07;  
 Matches 16; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKLNCRLYRKANKSSKL 17  
 |||||||||||||||  
 Db 125 AKLNCRLYRKANKSSKL 141

## RESULT 3

ANT3\_SHEEP STANDARD; PRT; 465 AA.  
 AC P32262;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-OCT-1993 (Rel. 27, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Antithrombin-III precursor (AIII).  
 GN SERPINC1 OR AT3.  
 OS Ovis aries (Sheep).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euthera; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;  
 OC Bovidae; Caprinae; Ovis.  
 OX NCBI\_TaxID=9940;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93129622; PubMed=1482684;  
 RA Niessen R.W.L.M., Sturk A., Hordijk P.L., Michiels F., Peters M.;  
 RT "Sequence characterization of a sheep cDNA for antithrombin III.";  
 RL Biochim. Biophys. Acta 1171:207-210(1992).  
 CC -!- FUNCTION: MOST IMPORTANT SERINE PROTEASE INHIBITOR IN PLASMA  
 CC THAT REGULATES THE BLOOD COAGULATION CASCADE. AT-III INHIBITS  
 CC THROMBIN AS WELL AS FACTORS IXA, XA AND XIA. ITS INHIBITORY  
 CC ACTIVITY IS GREATLY ENHANCED IN THE PRESENCE OF HEPARIN.  
 CC -!- SUBCELLULAR LOCATION: Extracellular.  
 CC -!- TISSUE SPECIFICITY: PLASMA.  
 CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

EMBL; X68287; CAA48347.1; -  
 DR PIR; S28219; S28219.  
 DR HSSP; P01008; IATH.  
 DR InterPro; IPR000215; Serpin.  
 DR Pfam; PF00079; serpin; 1.  
 DR SMART; SM00093; SERPIN; 1.  
 DR PROSITE; PS00284; SERPIN; 1.  
 KW Serpin; Serine protease inhibitor; Glycoprotein; Heparin-binding;  
 KW Plasma; Blood coagulation; Signal.  
 FT SIGNAL 1 32 BY SIMILARITY.  
 FT CHAIN 33 465 ANTITHROMBIN-III.  
 FT BINDING 82 82 HEPARIN (BY SIMILARITY).  
 FT BINDING 162 162 HEPARIN (BY SIMILARITY).  
 FT BINDING 178 178 HEPARIN (BY SIMILARITY).  
 FT ACT\_SITE 426 427 REACTIVE BOND.  
 FT DISULFID 41 161 BY SIMILARITY.  
 FT DISULFID 54 128 BY SIMILARITY.  
 FT DISULFID 280 463 BY SIMILARITY.  
 FT CARBOHYD 129 129 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 168 168 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 188 188 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 225 225 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 465 AA; 52499 MW; E075EFAE51D5F118 CRC64;

Query Match 95.3%; Score 82; DB 1; Length 465;  
 Best Local Similarity 94.1%; Pred. No. 1e-06;  
 Matches 16; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKLNCRLYRKANKSSKL 17  
 |||||||||||||||  
 Db 157 AKLNCRLYRKANKSSKL 173

## RESULT 4

ANT3\_MOUSE STANDARD; PRT; 465 AA.  
 ID ANT3\_MOUSE

AC P32261;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-OCT-1993 (Rel. 27, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Anthrombin-III precursor (ATIII).  
 GN SERPIN1 OR AT3.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A. PubMed=1440494;  
 RX MEDLINE=93069082; PubMed=1440494;  
 RA Wu J.K., Sheffield W.P., Blajchman M.A.;  
 RT "Molecular cloning and cell-free expression of mouse antithrombin  
 III.";  
 RL Thromb. Haemost. 68:291-296(1992).  
 CC -!- FUNCTION: MOST IMPORTANT SERINE PROTEASE INHIBITOR IN PLASMA  
 CC THAT REGULATES THE BLOOD COAGULATION CASCADE. AT-III INHIBITS  
 CC THROMBIN AS WELL AS FACTORS IXA, XA AND XIA. ITS INHIBITORY  
 CC ACTIVITY IS GREATLY ENHANCED IN THE PRESENCE OF HEPARIN.  
 CC -!- SUBCELLULAR LOCATION: Extracellular.  
 CC -!- TISSUE SPECIFICITY: PLASMA.  
 CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; S47225; AAB23965.1; -  
 CC HSPF; P01008; IATH.  
 DR MGD; MGI:88095; Serpincl.  
 DR InterPro: IPR000215; Serpin.  
 DR Pfam; PF00079; serpin; 1.  
 DR SMART; SM00093; serpin; 1.  
 DR PROSITE; PS00284; SERPIN; 1.  
 KW Serpin; Serine protease inhibitor; Glycoprotein; Heparin-binding;  
 KW Plasma; Blood coagulation; Signal.  
 FT SIGNAL 1 32  
 FT CHAIN 33 465  
 FT BINDING 82 82  
 FT BINDING 162 162  
 FT BINDING 178 178  
 FT ACT\_SITE 426 427  
 FT DISULFID 41 161  
 FT DISULFID 54 128  
 FT DISULFID 280 463  
 FT CARBOHYD 129 129  
 FT CARBOHYD 168 168  
 FT CARBOHYD 188 188  
 FT CARBOHYD 225 225  
 SQ SEQUENCE 465 AA; 52003 MW; 5CE087E98874E35D CRC64;  
 Query Match 93.0%; Score 80; DB 1; Length 465;  
 Best Local Similarity 94.1%; Pred. No. 2.2e-06;  
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 AKLCRLRYKANKSSKL 17  
 Db 157 AKLCRLRYKANKSSDL 173  
 RESULT 5  
 ACMS5\_RAT STANDARD; PRT; 531 AA.  
 AC P08911;  
 DT 01-NOV-1988 (Rel. 09, Created)  
 DT 01-NOV-1988 (Rel. 09, Last sequence update)  
 DT 01-NOV-1995 (Rel. 32, Last annotation update)

DE Muscarinic acetylcholine receptor M5.  
 GN CHRMS OR CHR5.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A. PubMed=3272174;  
 RX MEDLINE=90166521; PubMed=3272174;  
 RA Bonner T.I., Young A.C., Brann M.R., Buckley N.J.;  
 RT "Cloning and expression of the human and rat m5 muscarinic  
 RT acetylcholine receptor genes.";  
 RL Neuron 1:403-410(1988).  
 RN [2]  
 RP SEQUENCE FROM N.A. PubMed=2540186;  
 RX MEDLINE=89214170; PubMed=2540186;  
 RA Liao C.-F., Themmen A.P.N., Joho R., Barberis C., Birnbaumer M.,  
 RA Birnbaumer L.;  
 RT "Molecular cloning and expression of a fifth muscarinic acetylcholine  
 RT receptor.";  
 RL J. Biol. Chem. 264:7328-7337(1989).  
 CC -!- FUNCTION: THE MUSCARINIC ACETYLCHOLINE RECEPTOR MEDIATES VARIOUS  
 CC CELLULAR RESPONSES, INCLUDING INHIBITION OF ADENYLATE CYCLASE,  
 CC BREAKDOWN OF PHOSPHOINOSITIDES & MODULATION OF POTASSIUM CHANNELS  
 CC THROUGH THE ACTION OF G PROTEINS. PRIMARY TRANSDUCING EFFECT IS PI  
 CC TURNOVER.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; M22926; AAA40658.1; -  
 CC EMBL; M22925; AAA41572.1; -  
 CC PIR; JT0531; JT0531.  
 DR PIR; A33354; A33354.  
 DR InterPro: IPR000276; GPCR\_Rhodopsn.  
 DR Pfam; PF00001; 7tm\_1; 1.  
 DR PRINTS; PR00237; GPCR\_Rhodopsn.  
 DR PROSITE; PS00237; G-PROTEIN\_RECEP\_F1\_1; 1.  
 DR PROSITE; PS00262; G-PROTEIN\_RECEP\_F1\_2; 1.  
 KW Postsynaptic membrane; Ionic channel; Glycoprotein; Transmembrane;  
 KW Phosphorylation; Multigene family; G-protein coupled receptor.  
 FT DOMAIN 1 28  
 FT TRANSMEM 29 52  
 FT TRANSMEM 53 65  
 FT DOMAIN 56 86  
 FT TRANSMEM 87 103  
 FT DOMAIN 104 125  
 FT TRANSMEM 126 145  
 FT TRANSMEM 146 168  
 FT TRANSMEM 169 190  
 FT TRANSMEM 191 213  
 FT TRANSMEM 214 442  
 FT TRANSMEM 443 463  
 FT TRANSMEM 464 477  
 FT TRANSMEM 478 497  
 FT TRANSMEM 498 531  
 FT CARBOHYD 7 7  
 FT DISULFID 102 182  
 FT MOD\_RES 500 500  
 FT MOD\_RES 504 504  
 SQ SEQUENCE 531 AA; 60136 MW; 647CE0D5D75A2BB1 CRC64;  
 Query Match 47.7%; Score 41; DB 1; Length 531;  
 Best Local Similarity 50.0%; Pred. No. 13;  
 Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;



DR PROSITE; PS00211; ABC\_TRANSPORTER; 1.  
KW Hypothetical protein; ATP-binding; Transport; Transmembrane;  
KW Complete proteome.

```

FT TRANSMEM 27 47 POTENTIAL.
FT TRANSMEM 81 101 POTENTIAL.
FT TRANSMEM 157 177 POTENTIAL.
FT TRANSMEM 181 201 POTENTIAL.
FT TRANSMEM 266 286 POTENTIAL.
FT TRANSMEM 307 327 POTENTIAL.
FT NP_BIND 400 407 ATP (POTENTIAL).
FT CONFLICT 331 333 FNR -> LIV (IN REF. 2).
SQ SEQUENCE 623 AA; 69536 MW; F5D46215C9A595DDB CRC64;

Query Match 45.5%; Score 40; DB 1; Length 623;
Best Local Similarity 53.8%; Pred. No. 23;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 4 NCRLYRKANKSKK 16
 |||||:|
Db 601 NCSLYQKMKESQ 613

RESULT 10
ID YOFB_BACSU STANDARD; PRT; 285 AA.
AC O34685;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical transcriptional regulator yofA.
GN YOFB.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Azevedo F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Borris L., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA Enrian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Gallizzi A., Galleron N.,
RA Glim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Guisepi G., Guy B.-J., Haga K., Haiech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hulio M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Paro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tanakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;
FT "The complete genome sequence of the Gram-positive bacterium Bacillus
FT subtilis."
RL Nature 390:249-256(1997).
CC -!- SIMILARITY: BELONGS TO THE LYSR FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L00951; AAA29814.1; -
CC PIR; G44490; G44490.
CC InterPro; IPR000477; RVTse.
CC Pfam; PF000078; rvt; 1.
CC Transferase; RNA-directed DNA polymerase; Transposable element;
CC Hydrolase; Nuclease; Endonuclease.
CC NON_TER 1 1
CC FT DOMAIN <1 600 REVERSE TRANSCRIPTASE.
CC FT DOMAIN 601 869 NUCLEIC ACID-BINDING ENDONUCLEASE.
CC SQ SEQUENCE 869 AA; 97398 MW; 9584756BFA7AF818 CRC64;
Query Match 45.3%; Score 39; DB 1; Length 869;
Best Local Similarity 50.0%; Pred. No. 49;

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z99113; CAB13725.1; -
CC EMBL; Z99114; CAB13735.1; -
CC Subtilist; BG13552; YOFB.
CC InterPro; IPR000847; HTH_LYSR.
CC InterPro; IPR005119; LYSR_subst.
CC Pfam; PF00126; HTH_1; 1.
CC Pfam; PF03466; LYSR_substrate; 1.
CC PRINTS; PR00039; HTHLYSR.
CC PROSITE; PS00044; HTH_LYSR_FAMILY; 1.
CC Hypothetical protein; Transcription regulation; DNA-binding;
KW Complete proteome. 37 H-T-H MOTIF (POTENTIAL).
FT DNA_BIND 18
FT SEQUENCE 285 AA; 32050 MW; BB2E45D25673F73 CRC64;

Query Match 45.3%; Score 39; DB 1; Length 285;
Best Local Similarity 46.7%; Pred. No. 16;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 LNCRLYRKANKSKK 17
 |||||:|
Db 43 LNIRLFHRTNRGMKL 57

RESULT 11
ID PO21_SCICO STANDARD; PRT; 869 AA.
AC Q03279;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Retrovirus-related POL polyprotein from type I retrotransposable
DE element R2 [Contains: Reverse transcriptase (EC 2.7.7.49);
DE Endonuclease (Fragment)].
OS Sciara coprophila (Fungus gnat).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera;
OC Sciarioidea; Sciariidae; Bradysia.
OX NCBI_TaxID=38358;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93196484; PubMed=8383793;
RA Burke W.D., Eickbush D.G., Xiong Y., Jakubczak J.L., Eickbush T.H.;
RA "Sequence relationship of retrotransposable elements R1 and R2 within
RA and between divergent insect species."
RL Mol. Biol. Evol. 10:163-185(1993).
CC -!- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + (DNA)(N).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L00951; AAA29814.1; -
CC PIR; G44490; G44490.
CC InterPro; IPR000477; RVTse.
CC Pfam; PF000078; rvt; 1.
CC Transferase; RNA-directed DNA polymerase; Transposable element;
CC Hydrolase; Nuclease; Endonuclease.
CC NON_TER 1 1
CC FT DOMAIN <1 600 REVERSE TRANSCRIPTASE.
CC FT DOMAIN 601 869 NUCLEIC ACID-BINDING ENDONUCLEASE.
CC SQ SEQUENCE 869 AA; 97398 MW; 9584756BFA7AF818 CRC64;
Query Match 45.3%; Score 39; DB 1; Length 869;
Best Local Similarity 50.0%; Pred. No. 49;

```





```
FT CHAIN 28 944 LACTOFERRIN BINDING PROTEIN A.
FT SITE 827 944 TONB C-TERMINAL BOX.
SQ SEQUENCE 944 AA; 105732 MW; CE06B6192E74AE3E CRC64;

Query Match 45.3%; Score 39; DB 1; Length 944;
Best Local Similarity 46.7%; Pred. No. 53;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

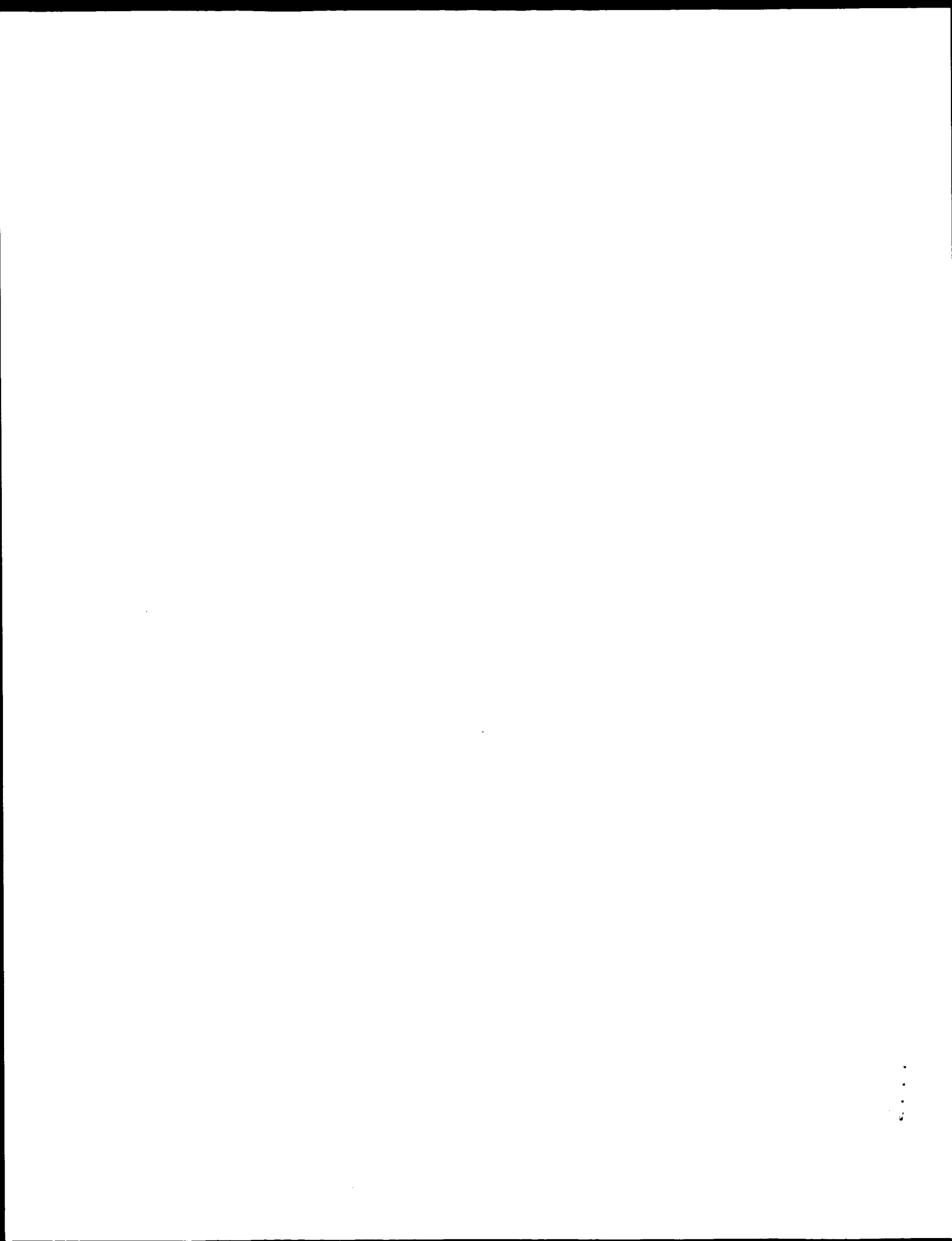
Qy 2 KLNCRLYRKANKSSK 16
Db 482 KLNCVAVYPAVDKSR 496

RESULT 14
BAR3_CHITE
ID BAR3_CHITE STANDARD; PRT; 1700 AA.
AC Q03376;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Balbiani ring protein 3 precursor.
GN BR3.
OS Chironomus tentans (Midge).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera;
OC Chironomidae; Chironomidae; Chironominae; Chironomus.
OX NCBI_TaxID=7153;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RX MEDLINE=90172404; PubMed=1689777;
RA Paulsson G., Lendahl U., Galli J., Ericsson C., Wieslander L.;
RT "The Balbiani ring 3 gene in Chironomus tentans has a diverged
RL J. Mol. Biol. 211:331-349(1990).
CC -!- FUNCTION: USED BY THE LARVAE TO CONSTRUCT A SUPRAMOLECULAR
CC STRUCTURE, THE LARVAL TUBE. BALBIANI RING PROTEIN 3 COULD PLAY A
CC ROLE AS A TRANSPORT PROTEIN THAT BINDS TO OTHER PROTEINS
CC INTRACELLULARLY AND IN THE GLAND LUMEN IN ORDER TO PREVENT THESE
CC FROM FORMING WATER-INSOLUBLE FIBERS TOO EARLY.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: SALIVARY GLAND.
CC -!- DOMAIN: HAS 82 APPROXIMATE REPEATS OF CYS-X-CYS-X-CYS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X52263; CAA36506.1; -
CC PIR: S08167; S08167.
CC HSP: P15358; ISK2.
CC InterPro: IPR004153; CXCXC_repeat.
CC Pfam: PF03128; CXCXC; 71.
CC Repeat: Signal.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 1700 BALBIANI RING PROTEIN 3.
SQ SEQUENCE 1700 AA; 186145 MW; 34202B28521B0815 CRC64;

Query Match 45.3%; Score 39; DB 1; Length 1700;
Best Local Similarity 53.3%; Pred. No. 97;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 AKLNCRLYRKANKSS 15
Db 275 AQLNCPDNKKANKET 289

RESULT 15
BIMB_EMENI
```



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 2, 2002, 10:00:15 ; Search time 6.08955 Seconds  
(without alignments)  
575.215 Million cell updates/sec

Title: US-09-741-106-12

Perfect score: 86

Sequence: 1 AKLNCRLYRKANKSSKL 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_21.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organelle.\*  
9: sp\_phase.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*  
15: sp\_virus.\*  
16: sp\_bacteriap.\*  
17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score | Query Match | Length | ID        | Description        |
|------------|-------|-------------|--------|-----------|--------------------|
| 1          | 86    | 100.0       | 259    | 4 Q8TCE1  | Q8tce1 homo sapien |
| 2          | 85    | 100.0       | 465    | 4 Q13815  | Q13815 homo sapien |
| 3          | 79    | 91.9        | 423    | 13 Q91422 | Q91422 gallus gall |
| 4          | 79    | 91.9        | 452    | 13 Q8UVX2 | Q8uvx2 sphendon p  |
| 5          | 79    | 91.9        | 453    | 13 Q8UVS2 | Q8uvs2 struthio ca |
| 6          | 79    | 91.9        | 453    | 13 Q8UVS1 | Q8uvs1 chelydra se |
| 7          | 78    | 90.7        | 456    | 13 Q8UVS0 | Q8uvs0 xenopus lae |
| 8          | 77    | 89.5        | 459    | 11 Q9WTT1 | Q9wtt1 cavia porce |
| 9          | 69    | 80.2        | 169    | 6 Q9GMD6  | Q9gmd6 sus scrofa  |
| 10         | 69    | 80.2        | 448    | 13 Q9W648 | Q9w648 fugu rubrip |
| 11         | 66    | 76.7        | 452    | 13 Q9PTA8 | Q9pta8 salmo salar |
| 12         | 46    | 53.5        | 302    | 16 Q987S1 | Q987s1 rhizobium 1 |
| 13         | 43    | 50.0        | 2010   | 5 Q96171  | Q96171 plasmodium  |
| 14         | 42    | 48.8        | 294    | 5 Q9N897  | Q9n897 plasmodium  |
| 15         | 42    | 48.8        | 407    | 16 Q34413 | Q34413 bacillus su |
| 16         | 42    | 48.8        | 713    | 10 Q9ASW9 | Q9asw9 arabidopsis |

|    |    |      |      |           |                     |
|----|----|------|------|-----------|---------------------|
| 17 | 42 | 48.8 | 716  | 10 Q8RXH2 | Q8rxh2 arabidopsis  |
| 18 | 41 | 47.7 | 62   | 10 Q9LUZ7 | Q9luz7 arabidopsis  |
| 19 | 41 | 47.7 | 285  | 16 Q911Y2 | Q911y2 pseudomonas  |
| 20 | 41 | 47.7 | 294  | 6 Q8WNM6  | Q8wnm6 gorilla gor  |
| 21 | 41 | 47.7 | 296  | 6 Q8WNM5  | Q8wnm5 pongo pygma  |
| 22 | 41 | 47.7 | 347  | 17 Q8TKP5 | Q8tkp5 methanosarc  |
| 23 | 41 | 47.7 | 498  | 15 P89686 | P89686 feline immu  |
| 24 | 41 | 47.7 | 530  | 11 Q8VH24 | Q8vvh24 cavia porce |
| 25 | 41 | 47.7 | 532  | 11 Q920H4 | Q920h4 mus musculu  |
| 26 | 41 | 47.7 | 651  | 12 Q9WNM8 | Q9wnm8 human papil  |
| 27 | 41 | 47.7 | 664  | 10 Q9STY9 | Q9sty9 arabidopsis  |
| 28 | 41 | 47.7 | 906  | 10 Q9LNF8 | Q9lfn8 arabidopsis  |
| 29 | 41 | 47.7 | 1287 | 4 Q92549  | Q92549 homo sapien  |
| 30 | 40 | 46.5 | 244  | 16 Q97EL1 | Q97ell clostridium  |
| 31 | 40 | 46.5 | 276  | 10 Q9MBC7 | Q9mbc7 oryza sativ  |
| 32 | 40 | 46.5 | 287  | 10 Q9ZRZ3 | Q9zrz3 triticum sp  |
| 33 | 40 | 46.5 | 331  | 5 Q97262  | Q97262 plasmodium   |
| 34 | 40 | 46.5 | 420  | 5 Q8SUY5  | Q8suy5 encephalito  |
| 35 | 40 | 46.5 | 560  | 13 Q8UVH4 | Q8uvh4 brachydanio  |
| 36 | 40 | 46.5 | 578  | 5 P90536  | P90536 dictyosteli  |
| 37 | 40 | 46.5 | 771  | 5 Q9VAQ1  | Q9vaq1 drosophila   |
| 38 | 40 | 46.5 | 2357 | 5 Q9UIM8  | Q9ulm8 dictyosteli  |
| 39 | 39 | 45.3 | 151  | 17 Q8TR99 | Q8trs9 methanosarc  |
| 40 | 39 | 45.3 | 206  | 5 Q9V8L1  | Q9v8l1 drosophila   |
| 41 | 39 | 45.3 | 213  | 5 Q9VBQ4  | Q9vbq4 drosophila   |
| 42 | 39 | 45.3 | 217  | 4 Q9BTQ1  | Q9btq1 homo sapien  |
| 43 | 39 | 45.3 | 223  | 5 Q8T1J1  | Q8t1j1 dictyosteli  |
| 44 | 39 | 45.3 | 234  | 7 Q91380  | Q91380 cyprinus ca  |
| 45 | 39 | 45.3 | 257  | 2 Q8VSO7  | Q8vs07 klebsiella   |

## ALIGNMENTS

### RESULT 1

Q8TCE1 PRELIMINARY; PRT; 259 AA.  
ID Q8TCE1  
AC Q8TCE1  
DT 01-JUN-2002 (TREMBlrel. 21, Created)  
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
DE Similar to serine (or cysteine) proteinase inhibitor, clade C  
DE (antithrombin), member 1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=LIVER;  
RA Straussberg R.;  
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC022309; AAH22309.1; -;  
SQ SEQUENCE 259 AA; 29092 MW; 8386705A82B8EC8D CRC64;

Query Match 100.0%; Score 86; DB 4; Length 259;  
Best Local Similarity 100.0%; Pred. No. 2.7e-07;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKLNCRLYRKANKSSKL 17

Db 156 AKLNCRLYRKANKSSKL 172

### RESULT 2

Q13815 PRELIMINARY; PRT; 465 AA.  
ID Q13815  
AC Q13815; P78447;  
DT 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
DT 01-OCT-2001 (TREMBlrel. 18, Last annotation update)  
DE Antithrombin III.  
GN AT3.

OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
[1]  
RP SEQUENCE FROM N.A.  
RA Tsuji H., Takada O., Nakagawa M., Tanaka S., Hashimoto-Gotoh T.;  
RL Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.  
[2]  
RP SEQUENCE OF 138-255 FROM N.A.  
RX MEDLINE=89050967; PubMed=3191114;  
RA Bock S.C., Marrinan J.A., Radziejewska E.;  
RT "Antithrombin III Utah: proline-407 to leucine mutation in a highly  
RL conserved region near the inhibitor reactive site";  
PL Biochemistry 27:6171-6178(1988).  
[3]  
RP SEQUENCE OF 1-14 FROM N.A.  
RX MEDLINE=84169500; PubMed=6672771;  
RA Bock S.C., Levitan D.J.;  
RT "Characterization of an unusual DNA length polymorphism 5' to the  
RL human antithrombin III gene";  
PL Nucleic Acids Res. 11:8569-8582(1983).  
[4]  
RP SEQUENCE OF 1-13 FROM N.A.  
RA Rosenberg J.B., Amrani D.L., Bergtrom G.;  
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.  
[5]  
RP SEQUENCE OF 1-13 FROM N.A.  
RX MEDLINE=85080031; PubMed=6096369;  
RA Prochownik E.V., Orkin S.H.;  
RT "In vivo transcription of a human antithrombin III 'minigene'";  
PL J. Biol. Chem. 259:15386-15392(1984).  
CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY.  
DR EMBL; M21644; AAA51794.1; -  
DR EMBL; M21643; AAA51794.1; JOINED.  
DR EMBL; M21643; AAA51793.1; -  
DR EMBL; D29832; BAA06212.1; -  
DR EMBL; X00238; CAA25060.1; -  
DR EMBL; X00237; CAA25059.1; -  
DR EMBL; U11270; AAA19930.1; -  
DR HSP; P01008; 2AANT  
DR InterPro; IPR000215; Serpin.  
DR Pfam; PF000079; serpin; 1.  
DR SMART; SM00093; SERPIN; 1.  
KW Serpin; Signal.  
SQ SEQUENCE 465 AA; 52691 MW; C69214FCA12766D0 CRC64;  
Query Match 100.0%; Score 86; DB 4; Length 465;  
Best Local Similarity 100.0%; Pred. No. 4.8e-07;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AKLNCRLYRKANKSKSL 17  
DB 157 AKLNCRLYRKANKSKSL 173  
RESULT 3  
Q91422  
ID Q91422 PRELIMINARY; PRT; 423 AA.  
AC Q91422;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Antithrombin (Fragment).  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96114615; PubMed=7495074;  
RA Tejada M.L., Degley R.G.;

RT "Cloning of an avian antithrombin: developmental and hormonal  
RL regulation of expression";  
RL Thromb. Haemost. 73:654-661(1995).  
CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY.  
DR EMBL; S79838; AAB35653.1; -  
DR HSP; P41361; IAIT.  
DR InterPro; IPR000215; Serpin.  
DR Pfam; PF00079; serpin; 1.  
DR SMART; SM00093; SERPIN; 1.  
DR PROSITE; PS00284; SERPIN; UNKNOWN\_1.  
KW Serpin.  
SQ SEQUENCE 423 AA; 48128 MW; 5C559B91312D1F25 CRC64;  
FT NON-TER 1  
Query Match 91.9%; Score 79; DB 13; Length 423;  
Best Local Similarity 88.2%; Pred. No. 7.7e-06;  
Matches 15; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AKLNCRLYRKANKSKSL 17  
DB 115 AKLNCRLYRKANKSKSEL 131  
RESULT 4  
Q8UVX2  
ID Q8UVX2 PRELIMINARY; PRT; 452 AA.  
AC Q8UVX2;  
DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE Antithrombin III.  
GN ATIII.  
OS Sphenodon punctatus (Hatteria) (Tuatara).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Lepidosauria; Sphenodontia; Sphenodontidae; Sphenodon.  
OX NCBI\_TaxID=8508;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Metcalf V.J., Brennan S.O., George P.M., Chambers G.K.;  
RT "The identification of liver-expressed genes from tuatara";  
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF335570; AAL73207.1; -  
DR InterPro; IPR000215; Serpin.  
DR Pfam; PF00079; serpin; 1.  
DR SMART; SM00093; SERPIN; 1.  
DR PROSITE; PS00284; SERPIN; UNKNOWN\_1.  
SQ SEQUENCE 452 AA; 51807 MW; 380C1C653959DA07 CRC64;  
Query Match 91.9%; Score 79; DB 13; Length 452;  
Best Local Similarity 88.2%; Pred. No. 8.2e-06;  
Matches 15; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AKLNCRLYRKANKSKSL 17  
DB 144 AKLNCRLYRKANKSKSEL 160  
RESULT 5  
Q8UVS2  
ID Q8UVS2 PRELIMINARY; PRT; 453 AA.  
AC Q8UVS2;  
DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE Antithrombin.  
OS Struthio camelus (Ostrich).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Palaeognathae; Struthioniformes; Struthionidae;  
OC Struthio.  
OX NCBI\_TaxID=8801;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Backovic M., Gettins P.G.W.;

RT "Insight into residues critical for antithrombin function from an  
RT expanded database of sequences that includes frog, turtle and ostrich  
RT antithrombins.";  
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF411691; AAL60465.1; -.  
DR InterPro; IPR000215; Serpin.  
DR Pfam; PF00079; serpin; 1.  
DR SMART; SM00093; SERPIN; 1.  
DR PROSITE; PS00284; SERPIN; UNKNOWN\_1.  
SQ SEQUENCE 453 AA; 51403 MW; 4283370BF0C6CFFF CRC64;

Query Match 91.9%; Score 79; DB 13; Length 453;  
Best Local Similarity 88.2%; Pred. No. 8.2e-06;  
Matches 15; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKLNCRLYRKANKSSKL 17

|||||:|||||:|

Db 145 AKLNCRLYRKANKSSSEL 161

RESULT 6

ID Q8UVS1 PRELIMINARY; PRT; 453 AA.

AC Q8UVS1; 20, Created)

DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)

DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE Antithrombin.

OS Chelydra serpentina (Snapping turtle).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Testudines; Cryptodira; Testudinoidea; Chelydridae; Chelydra.

OX NCBI\_TaxID=8475;

RN [1]

RP SEQUENCE FROM N.A.

RA Backovic M., Gettins P.G.W.;

RT "Insight into residues critical for antithrombin function from an  
RT expanded database of sequences that includes frog, turtle and ostrich  
RT antithrombins.";

RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF411692; AAL60466.1; -.  
DR InterPro; IPR000215; Serpin.  
DR Pfam; PF00079; serpin; 1.  
DR SMART; SM00093; SERPIN; 1.  
DR PROSITE; PS00284; SERPIN; UNKNOWN\_1.  
SQ SEQUENCE 453 AA; 51593 MW; A86A8ABC7798799E CRC64;

Query Match 91.9%; Score 79; DB 13; Length 453;

Best Local Similarity 88.2%; Pred. No. 8.2e-06;

Matches 15; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKLNCRLYRKANKSSKL 17

|||||:|||||:|

Db 145 AKLNCRLYRKANKSSSEL 161

RESULT 7

ID Q8UVS0 PRELIMINARY; PRT; 456 AA.

AC Q8UVS0; 20, Created)

DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)

DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE Antithrombin.

OS Xenopus laevis (African clawed frog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;

OC Xenopodinae; Xenopus.

OX NCBI\_TaxID=6355;

RN [1]

RP SEQUENCE FROM N.A.

RA Backovic M., Gettins P.G.W.;

RT "Insight into residues critical for antithrombin function from an  
RT expanded database of sequences that includes frog, turtle and ostrich

RT antithrombins.";  
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF411693; AAL60467.1; -.  
DR InterPro; IPR000215; Serpin.  
DR Pfam; PF00079; serpin; 1.  
DR SMART; SM00093; SERPIN; 1.  
DR PROSITE; PS00284; SERPIN; UNKNOWN\_1.  
SQ SEQUENCE 456 AA; 51851 MW; 03E086F3ABCEA85C CRC64;

Query Match 90.7%; Score 78; DB 13; Length 456;

Best Local Similarity 88.2%; Pred. No. 1.2e-05;

Matches 15; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKLNCRLYRKANKSSKL 17

|||||:|||||:|

Db 147 AKLNCRLYRKANKSSSEL 163

RESULT 8

ID Q9WTT1 PRELIMINARY; PRT; 459 AA.

AC Q9WTT1; 12, Created)

DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)

DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE Antithrombin III.

OS Cavia porcellus (Guinea pig).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Hystriognathi; Cavidae; Cavia.

OX NCBI\_TaxID=10141;

RN [1]

RP SEQUENCE FROM N.A.

RA Suzuki Y.;

RT "Guinea pig gene for antithrombin III.";

RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.

CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY.

DR EMBL; AB027238; BAA77781.1; -.  
DR HSP; P01008; 1ATH.  
DR InterPro; IPR000215; Serpin.  
DR Pfam; PF00079; serpin; 1.  
DR SMART; SM00093; SERPIN; 1.  
DR PROSITE; PS00284; SERPIN; UNKNOWN\_1.  
KW Serpin.  
SQ SEQUENCE 459 AA; 51912 MW; B1E36FB732592824 CRC64;

Query Match 89.5%; Score 77; DB 11; Length 459;

Best Local Similarity 100.0%; Pred. No. 1.9e-05;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKLNCRLYRKANKSS 15

|||||:|||||:|

Db 151 AKLNCRLYRKANKSS 165

RESULT 9

ID Q9GMD6 PRELIMINARY; PRT; 169 AA.

AC Q9GMD6; 16, Created)

DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)

DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE Antithrombin III (Fragment).

GN AT3.

OS Sus scrofa (Pig).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

OX NCBI\_TaxID=9823;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=LIVER;

RA Schuetz E., Andag R., Wieland E.;

RT "Landrace pig Antithrombin 3, partial cDNA.";

RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.

```
DR EMBL; AF281653; AAG01450.1; -.
DR HSSP; P41361; IATH.
DR InterPro; IPR000215; Serpin.
DR Pfam; PF00079; serpin; 1.
FT NON_TER 169
SQ SEQUENCE 169 AA; 19170 MW; 182B1BC2DIC11A21 CRC64;

Query Match 80.2%; Score 69; DB 6; Length 169;
Best Local Similarity 100.0%; Pred. No. 0.00019;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKLNCRLYRKANK 13
| | | | | | | | | | | | | |
Db 157 AKLNCRLYRKANK 169

RESULT 10
Q9W648 PRELIMINARY; PRT; 448 AA.
AC Q9W648;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Anthrombin III.
OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Takifugu.
OX NCBI_TaxID=31033;
RN [1]
RP SEQUENCE FROM N.A.
RA Suzuki Y.;
RT "pufferfish gene for antithrombin III.";
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
DR EMBL; AB026832; BAAT7461.1; -.
DR HSSP; P01008; IATH.
DR InterPro; IPR000215; Serpin.
DR Pfam; PF00079; serpin; 1.
DR SMART; SM00093; SERPIN; 1.
DR PROSITE; PS00284; SERPIN; UNKNOWN_1.
KW Serpin.
SQ SEQUENCE 448 AA; 50585 MW; A9037E2FCD5D3FE CRC64;

Query Match 80.2%; Score 69; DB 13; Length 448;
Best Local Similarity 76.5%; Pred. No. 0.00048;
Matches 13; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AKLNCRLYRKANKSKL 17
| | | | | | | | | | | | | |
Db 139 AKLNCRLYRKDKSNEL 155

RESULT 11
Q9PTA8 PRELIMINARY; PRT; 452 AA.
AC Q9PTA8;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Anthrombin precursor.
GN ANTHITHROMBIN.
OS Salmo salar (Atlantic salmon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
OX NCBI_TaxID=8030;
RN [1]
RP SEQUENCE FROM N.A.
RA Andersen O., Flengsrud R., Norberg K., Salte R.;
RX MEDLINE=20177689; PubMed=10712595;
RT "Salmon antithrombin has only three carbohydrate side chains, and
```

```
RT shows functional similarities to human beta-antithrombin.";
RL Eur. J. Biochem. 267:1651-1657(2000).
CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
DR EMBL; AJ252153; CAB64714.1; -.
DR HSSP; P01008; IATH.
DR InterPro; IPR000215; Serpin.
DR Pfam; PF00079; serpin; 1.
DR SMART; SM00093; SERPIN; 1.
DR PROSITE; PS00284; SERPIN; UNKNOWN_1.
KW Serpin; Signal.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 452 ANTHITHROMBIN.
SQ SEQUENCE 452 AA; 51471 MW; 797568B072B46E8 CRC64;

Query Match 76.7%; Score 66; DB 13; Length 452;
Best Local Similarity 70.6%; Pred. No. 0.0017;
Matches 12; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AKLNCRLYRKANKSKL 17
| | | | | | | | | | | | | |
Db 143 AKLNCRLYRKDKKTTEL 159

RESULT 12
Q987S1 PRELIMINARY; PRT; 302 AA.
AC Q987S1;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Transcriptional regulator.
GN MLL6938.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF30309;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti.";
RL DNA Res. 7:331-338(2000).
DR EMBL; AP003010; BAB53129.1; -.
DR InterPro; IPR000847; HTH_LysR.
DR InterPro; IPR005119; LysR_subst.
DR Pfam; PF00126; HTH_1; 1.
DR Pfam; PF03466; LysR_substrate; 1.
DR PRINTS; PR00039; HTHLYSR.
DR PROSITE; PS00044; HTH_LYSR_FAMILY; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 302 AA; 32873 MW; 630BE56B26FEC4D5 CRC64;

Query Match 53.5%; Score 46; DB 16; Length 302;
Best Local Similarity 43.8%; Pred. No. 3.9;
Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Qy 2 KLNCRLYRKANKSKL 17
| | | | | | | | | | | | | |
Db 45 KLNCRLFERGGKSVRM 60

RESULT 13
Q96171 PRELIMINARY; PRT; 2010 AA.
ID Q96171
AC Q96171;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
```

```

DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
DE Phosphatase (acid phosphatase family).
GN PFB0380C.
OS Plasmodium falci-parum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
[1]
RN SEQUENCE FROM N.A.
RX MEDLINE=99021743; PubMed=9804551;
RA Gardner M.J., Tetelin H., Carucci D.J., Cummings L.M., Aravind L.,
RA Koonin E.V., Shallow S., Mason T., Yu K., Fujii C., Pederson J.,
RA Shen K., Jing J., Aston C., Lai Z., Schwartz D.C., Pertea M.,
RA Salzberg S., Zhou L., Sutton G.G., Clayton R., White O., Smith H.O.,
RA Fraser C.M., Adams M.D., Venter J.C., Hoffman S.L.;
RT "Chromosome 2 sequence of the human malaria parasite Plasmodium
RT falci-parum.";
RL Science 282:1126-1132(1998).
DR EMBL; AE001391; AAC7185.1; -.
DR InterPro; IP0001064; Crystallin.
DR PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; UNKNOWN_1.
SQ SEQUENCE 2010 AA; 241791 MW; 6E30760AB70AA2F3 CRC64;

Query Match 50.0%; Score 43; DB 5; Length 2010;
Best Local Similarity 57.1%; Pred. No. 85;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 LNCRLYRKANKSSK 16
Db 1359 INVLRYEYKRNK 1372
:::|||||:::

RESULT 14
Q9N897 PRELIMINARY; PRT; 294 AA.
AC Q9N897;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)
DE Vir9 protein.
DE VIR9.
OS Plasmodium vivax.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5855;
[1]
RN SEQUENCE FROM N.A.
RA Oliver K., Bowman S., Hall N., Quail M., Rajandream M.A., Harris D.,
RA del Portillo H.A., Lanzer M., Barrell B.G.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL360334; CAB96698.1; -.
SQ SEQUENCE 294 AA; 34499 MW; 54F05C523996ED57 CRC64;

Query Match 48.8%; Score 42; DB 5; Length 294;
Best Local Similarity 53.3%; Pred. No. 20;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 3 LNCRLYRKANKSSK 17
Db 248 LSLRMVKKNKYSNI 262
I:|:|||||:

RESULT 15
O34413 PRELIMINARY; PRT; 407 AA.
AC O34413;
DT 01-JAN-1998 (TRENBLrel. 05, Created)
DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)
DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
DE YTCG.
DE YTCG.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;

```

```

RN SEQUENCE FROM N.A.
RP Lapidus A., Galleron N., Sorokin A., Ehrlich D.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE FROM N.A.
RP STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moser I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borriess R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA Etian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Funa S., Galizzi A., Galleron N.,
RA Chim S.Y., Glaser P., Goffeau A., Golligly E.J., Grandi G.,
RA Giuseppe G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter K., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler H., Wedler H., Weitzenecker T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the gram-positive bacterium Bacillus
RT subtilis.";
RL Nature 390:249-256(1997).
[3]
RN SEQUENCE FROM N.A.
RP STRAIN=168;
RC Kunst F., Ogasawara N., Yoshikawa H., Danchin A.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF008220; AAC00365.1; -.
DR EMBL; Z99119; CAB15066.1; -.
DR InterPro; IPR001296; Glycos_transf_1.
DR Pfam; PF00534; Glycos_transf_1; 1.
KW Complete proteome.
SQ SEQUENCE 407 AA; 46160 MW; BB6C767461A136D2 CRC64;

Query Match 48.8%; Score 42; DB 16; Length 407;
Best Local Similarity 53.3%; Pred. No. 27;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 3 LNCRLYRKANKSSK 17
Db 90 LNVSLYKASPNSKI 104
I:|:|||||:

Search completed: December 2, 2002, 10:04:12
Job time : 8.08955 secs

```

